

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:19:16 ; Search time 86 Seconds
(without alignments)
804.707 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MKRPYVAIVGRPNVCKSTIF.....IRAFGEFGTPIHIARKRN 436

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2213	100.0	22	AAU37007
2	2213	100.0	22	AAU37139
3	2213	100.0	22	AAU37139
4	2207	99.7	22	AAU37139
5	2094	94.6	23	ABP38338
6	1731	78.2	23	ABP38338
7	1528	69.0	23	ABP27827
8	1524	68.9	23	ABP27827
9	1523	68.8	23	ABP27826

10	1522	68.8	436	21	AAU04108	lyphc protein of st
11	1522	68.8	436	21	AAU04108	Streptococcus pneu
12	1522	68.8	436	24	ABU02185	S. pneumoniae type
13	1520	68.7	436	23	ABU04058	Lactococcus lactis
14	993.5	44.9	448	22	AAU09939	ERA binding domain
15	935.5	42.3	448	22	AAU09941	ERA binding domain
16	923	41.7	503	22	AAU34658	E. coli cellular p
17	907	41.0	501	22	AAU34658	ERA binding domain
18	899	40.6	519	22	AAU38110	Salmonella typhi c
19	862.5	39.0	519	22	AAU38110	C glutamic prote
20	857.5	38.7	500	22	ABP78407	N. gonorrhoeae aml
21	852	38.5	504	22	AAU35376	Haemophilus influe
22	852	38.5	504	22	AAU35376	Haemophilus influe
23	852	38.5	504	23	AAU1457	Haemophilus influe
24	851.5	38.5	709	23	ABP65695	Bifidobacterium lo
25	845.5	38.2	461	22	AAU0718	Propionibacterium
26	841.5	38.0	246	22	AAU0718	Streptococcus pneu
27	840.5	38.0	493	22	AAU06332	Pseudomonas aerugi
28	835	37.7	502	22	AAU09943	ERA binding domain
29	799.5	36.1	459	22	AAU09946	ERA binding domain
30	776.5	35.1	225	22	AAU09945	ERA binding domain
31	754.5	34.1	225	22	AAU033525	Enterococcus faeca
32	750.5	33.9	225	22	AAU03506	Enterococcus faeca
33	708.5	32.0	493	20	AAU37461	Amino acid sequenc
34	707	31.9	456	22	AAU09937	Helicobacter pylor
35	706	31.9	458	22	AAU35777	Helicobacter pylor
36	685.5	31.0	487	20	AAU35489	Amino acid sequenc
37	669.5	30.3	453	22	AAU09944	ERA binding domain
38	624	28.2	433	22	AAU09948	ERA binding domain
39	620.5	28.0	530	22	AAU09947	ERA binding domain
40	613	27.7	292	22	AAU09940	ERA binding domain
41	559.5	25.3	383	17	AAU77678	Campylobacter jelu
42	528.5	23.9	331	23	ABU51840	Helicobacter pylor
43	433.5	19.6	208	22	AAU09938	ERA binding domain
44	340	15.4	106	22	AAU09939	ERA binding domain
45	315	14.2	92	19	AAU86046	S. pneumoniae deri

ALIGNMENTS

RESULT 1	AAU37007	standard; Protein: 436 AA.
ID	AAU37007	
XX	AC	AAU37007:
XX	DT	14-FEB-2002 (first entry)
XX	DE	Staphylococcus aureus cellular proliferation protein #1177.
XX	DE	Staphylococcus aureus cellular proliferation protein #1177.
XX	KW	Antisense: prokaryotic cellular proliferation protein:
XX	KW	antibiotic; antibacterial; drug design.
XX	OS	Staphylococcus aureus.
XX	PN	W0200170955-A2.
XX	PD	27-SEP-2001.
XX	PF	21-MAR-2001; 2001WO-US09180.
XX	PR	21-MAR-2000; 2000US-191078P.
XX	PR	23-MAY-2000; 2000US-206848P.
XX	PR	26-MAY-2000; 2000US-207727P.
XX	PR	23-OCT-2000; 2000US-242578P.
XX	PR	27-NOV-2000; 2000US-253625P.
XX	PR	22-DEC-2000; 2000US-257931P.
XX	PR	16-FEB-2001; 2001US-269308P.
XX	PA	(ELIT-) ELITRA PHARM INC.
XX	PI	Hasselbeck R, Ohlsen KL, Zyckind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR N-PSDB: AAS54866.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12600; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 436 AA:
Query Match 100.0%; Score 2213; DB 22; Length 436;
Best Local Similarity 100.0%; Pred. No. 4,7e-184;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKPIVAIVGRPNNGKSTIFNRIVGERYSIYEDTPGVTDRDRYSSGEWLTDFDNITIDGG 60
DB 1 MTKPIVAIVGRPNNGKSTIFNRIVGERYSIYEDTPGVTDRDRYSSGEWLTDFDNITIDGG 60
QY 61 IEIGDAPFQTOIRAOAEIAIDEADVIIFMVNVRREGLTQSDENVAAQILKSKKPPVLA VNK 120
DB 61 IEIGDAPFQTOIRAOAEIAIDEADVIIFMVNVRREGLTQSDENVAAQILKSKKPPVLA VNK 120
QY 121 VDNMEMRTDVIYDFYSLGGEPPYPISSGSHGLGLDLDAAVSHFGEEDDPYDEDTIRLSI 180
DB 121 VDNMEMRTDVIYDFYSLGGEPPYPISSGSHGLGLDLDAAVSHFGEEDDPYDEDTIRLSI 180
QY 181 IGRNVGKSSLVNAALGDRIVSNVAGTTRDAIDTEYSYGDQYVLLDTAGMRKKGVY 240
DB 181 IGRNVGKSSLVNAALGDRIVSNVAGTTRDAIDTEYSYGDQYVLLDTAGMRKKGVY 240
QY 241 ESTEKYSVLRALKAIERSNVLVYDAEOGIIEODKRVAGAHGEGKAVVIVNKMWDVE 300
DB 241 ESTEKYSVLRALKAIERSNVLVYDAEOGIIEODKRVAGAHGEGKAVVIVNKMWDVE 300
QY 241 ESTEKYSVLRALKAIERSNVLVYDAEOGIIEODKRVAGAHGEGKAVVIVNKMWDVE 300
DB 241 ESTEKYSVLRALKAIERSNVLVYDAEOGIIEODKRVAGAHGEGKAVVIVNKMWDVE 300
QY 301 KDSKTMKKFEDEVKKEFOFLDYAOIAFVSAKERTRLRTLPYINDESENHKKRVOSSTLN 360
DB 301 KDSKTMKKFEDEVKKEFOFLDYAOIAFVSAKERTRLRTLPYINDESENHKKRVOSSTLN 360
QY 361 EYVVDATSMNPTPPDKGRLLNVFAATQVAIEPPTFFVNVNDELHMSYKRYLENOITAA 420
DB 361 EYVVDATSMNPTPPDKGRLLNVFAATQVAIEPPTFFVNVNDELHMSYKRYLENOITAA 420
QY 421 FGEGTPIHITARRKN 436
DB 421 FGEGTPIHITARRKN 436
RESULT 2
AAU37139
ID AAU37139 standard; Protein: 436 AA.
XX

AC AAU37139;
XX
XX 14-FEB-2002 (first entry)
DT
XX Staphylococcus aureus cellular proliferation protein #1309.
DE
XX
XX Antisense: prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-0S09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI: 2001-611495/70.
DR N-PSDB: AAS54998.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX
XX Example 3; Seq ID No 12732; 511pp; English.
PS
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 436 AA:
Query Match 100.0%; Score 2213; DB 22; Length 436;
Best Local Similarity 100.0%; Pred. No. 4,7e-184;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKPIVAIVGRPNNGKSTIFNRIVGERYSIYEDTPGVTDRDRYSSGEWLTDFDNITIDGG 60
DB 1 MTKPIVAIVGRPNNGKSTIFNRIVGERYSIYEDTPGVTDRDRYSSGEWLTDFDNITIDGG 60
QY 61 IEIGDAPFQTOIRAOAEIAIDEADVIIFMVNVRREGLTQSDENVAAQILKSKKPPVLA VNK 120
DB 61 IEIGDAPFQTOIRAOAEIAIDEADVIIFMVNVRREGLTQSDENVAAQILKSKKPPVLA VNK 120
QY 121 VDNMEMRTDVIYDFYSLGGEPPYPISSGSHGLGLDLDAAVSHFGEEDDPYDEDTIRLSI 180
DB 121 VDNMEMRTDVIYDFYSLGGEPPYPISSGSHGLGLDLDAAVSHFGEEDDPYDEDTIRLSI 180

|||||
Db 121 VDNMEMRTDYYDFYSLGFGPEPPIGSHGLGLDLDLDAVVSHFGEEDPYDEDTRLISI 180
QY 181 IGRPNVGKSSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGVY 240
Db 181 IGRPNVGKSSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGVY 240
QY 241 ESTEKYSYLRALKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGEGKAVIYVNMWDIVE 300
Db 241 ESTEKYSYLRALKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGEGKAVIYVNMWDIVE 300
QY 301 KDSKTMKKFEDEVKKEFOFLDYAOIAFVSASAKERTRLTLFPYINEASENHKKRVQSTLN 360
Db 301 KDSKTMKKFEDEVKKEFOFLDYAOIAFVSASAKERTRLTLFPYINEASENHKKRVQSTLN 360
QY 361 EYVTDATISMNPPTDQGRRLNVEYATQVAIEPPTFVFVNDVLMHFSYKRLLENQIRAA 420
Db 361 EYVTDATISMNPPTDQGRRLNVEYATQVAIEPPTFVFVNDVLMHFSYKRLLENQIRAA 420
QY 421 FGFEETPIHIIARKRN 436
Db 421 FGFEETPIHIIARKRN 436

RESULT 3
AAB82089
ID AAB82089 standard; Protein: 436 AA.

AC AAB82089;
DT 26-JUN-2001 (first entry)

Staphylococcus aureus Yphc protein sequence.

Yphc: antimicrobial; cytostatic; antiulcer; microbial infection;
gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
gastritis.

Staphylococcus aureus.

WO200123418-A1.

05-APR-2001.

19-SEP-2000; 2000WO-US25566.

28-SEP-1999; 99US-0406968.

(SMIK) SMITHKLINE BEECHAM CORP.
(SMIK) SMITHKLINE BEECHAM PLC.

Zalacain M, Biswas S, Burnham MKR, Sylvester D, Mcdevitt D;
Matile TB;

WPI; 2001-308138/32.
N-PSDB; AAF86461.

Novel yphc polypeptides of Staphylococcus aureus useful for diagnosing
and treating microbial infections, especially infection by
Staphylococcus aureus and Helicobacter pylori

Claim 1; Page 3; 41pp; English.

The present sequence is the yphc protein of Staphylococcus aureus. The
yphc coding sequence and protein are useful for treating and diagnosing
microbial infections such as infection caused by S.aureus and
Helicobacter pylori. In addition, the yphc coding sequence and protein
are useful for treating diseases such as H.pylori-induced cancers, e.g.
gastrointestinal carcinoma, gastric ulcers, and gastritis. The present
sequence was obtained from a library of clones of chromosomal DNA of
S.aureus in E.coli. The sequencing data from two or more clones
comprising overlapping S.aureus DNAs was used to construct the present
contiguous DNA sequence.

XX SQ Sequence 436 AA;
Query Match 100.0%; Score 2213; DB 22; Length 436;
Best Local Similarity 100.0%; Pred. No. 4.7e-184;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKPIVATIGRPNVGKSTIFNRIYGERVIVSDTFCGVTDRISYSGEMTLHPFNIIIDTG 60
Db 1 MTKPIVATIGRPNVGKSTIFNRIYGERVIVSDTFCGVTDRISYSGEMTLHPFNIIIDTG 60
QY 61 IEIGDAPFQTOIRAQAEIAIDEADVIFMVNVRGLTQSDENVAOILYKSKRPVLAUNK 120
Db 61 IEIGDAPFQTOIRAQAEIAIDEADVIFMVNVRGLTQSDENVAOILYKSKRPVLAUNK 120
QY 121 VDNMEMRTDYYDFYSLGFGPEPPIGSHGLGLDLDLDAVVSHFGEEDPYDEDTRLISI 180
Db 121 VDNMEMRTDYYDFYSLGFGPEPPIGSHGLGLDLDLDAVVSHFGEEDPYDEDTRLISI 180
QY 181 IGRPNVGKSSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGVY 240
Db 181 IGRPNVGKSSLVNALIGEDRVIVSNVAGTTRRAIDTEYSYDQDQVLLIDTAGMRKKGVY 240
QY 241 ESTEKYSYLRALKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGEGKAVIYVNMWDIVE 300
Db 241 ESTEKYSYLRALKAIERSNVVLVIDAEGGIIEDQKRVAGYAHGEGKAVIYVNMWDIVE 300
QY 301 KDSKTMKKFEDEVKKEFOFLDYAOIAFVSASAKERTRLTLFPYINEASENHKKRVQSTLN 360
Db 301 KDSKTMKKFEDEVKKEFOFLDYAOIAFVSASAKERTRLTLFPYINEASENHKKRVQSTLN 360
QY 361 EYVTDATISMNPPTDQGRRLNVEYATQVAIEPPTFVFVNDVLMHFSYKRLLENQIRAA 420
Db 361 EYVTDATISMNPPTDQGRRLNVEYATQVAIEPPTFVFVNDVLMHFSYKRLLENQIRAA 420
QY 421 FGFEETPIHIIARKRN 436
Db 421 FGFEETPIHIIARKRN 436
RESULT 4
AAU33787
ID AAU33787 standard; Protein: 435 AA.
AC AAU33787;
DT 14-FEB-2002 (first entry)
DE Staphylococcus aureus cellular proliferation protein #63.
KW Antisense; prokaryotic cellular proliferation protein;
antibiotic; antibacterial; drug design.
OS Staphylococcus aureus.
PN WO200170955-A2.
PD 27-SEP-2001.
PF 21-MAR-2001; 2001WO-US09180.
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
PA (ELIT-) ELITRA PHARM INC.
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.
DR N-PSDB: AAS51646.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
PS Example 3; Seq ID No 5283; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 435 AA:

Query Match 99.7%; Score 2207; DB 22; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.6e-183;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTHTDNIITDGG 60
DB 1 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTHTDNIITDGG 60
QY 1 IETGDAPFQOTIRAOAEIAIDADVITIFMVNREGLTQSDENVAQILYKSKRPVLAIVNK 120
DB 61 IETGDAPFQOTIRAOAEIAIDADVITIFMVNREGLTQSDENVAQILYKSKRPVLAIVNK 120
QY 121 VDNMEKRTDYYDFSYSGFGEPPYISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 180
DB 121 VDNMEKRTDYYDFSYSGFGEPPYISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 180
QY 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTRDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 240
DB 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTRDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 240
QY 241 ESTEKYSVLRAKAIERSNVVLVDAEGLIEQDKRVAGYAHEOGKAVVIVNKKMDTYE 300
DB 241 ESTEKYSVLRAKAIERSNVVLVDAEGLIEQDKRVAGYAHEOGKAVVIVNKKMDTYE 300
QY 301 KOSKTKKKFEDEVRKKEFOFLDYAQIAFVSAKERTRLTLTPYINSESENHKKRVOSSTLN 360
DB 301 KOSKTKKKFEDEVRKKEFOFLDYAQIAFVSAKERTRLTLTPYINSESENHKKRVOSSTLN 360
QY 361 EYVTDIAISMNPPTDGRRLNVFYATQVAIEPPTFVAVNDVLMHFSYRYLENOIRAA 420
DB 361 EYVTDIAISMNPPTDGRRLNVFYATQVAIEPPTFVAVNDVLMHFSYRYLENOIRAA 420
QY 421 FGEFGTPIHIIARKR 435
DB 421 FGEFGTPIHIIARKR 435

RESULT 5
ABP38338 standard; Protein; 443 AA.
ID ABP38338
XX AC ABP38338.

XX 24-JUL-2002 (first entry)
XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3183.
DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX Staphylococcus epidermidis; gene therapy.
XX Staphylococcus epidermidis.
OS
XX US6380370-B1.
PN 30-APR-2002.
XX 13-AUG-1998; 98US-0134001.
XX 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Doucette-Stamm LA, Bush D;
PI
XX WPI: 2002-381255/41.
DR N-PSDB: ABN90883.
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 3183; 267pp; English.



XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 443 AA:

Query Match 94.6%; Score 2094; DB 23; Length 443;
Best Local Similarity 93.1%; Pred. No. 1.1e-173;
Matches 406; Conservative 18; Mismatches 12; Indels 0; Gaps 0;
QY 1 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTHTDNIITDGG 60
DB 8 MKPPIAIVGRPNVSGSTIFNRIVGERVSIVEDTPGVTDRIRYSSGEMLTHTDNIITDGG 67
QY 61 IETGDAPFQOTIRAOAEIAIDADVITIFMVNREGLTQSDENVAQILYKSKRPVLAIVNK 120
DB 61 IETGDAPFQOTIRAOAEIAIDADVITIFMVNREGLTQSDENVAQILYKSKRPVLAIVNK 127
QY 121 VDNMEKRTDYYDFSYSGFGEPPYISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 180
DB 121 VDNMEKRTDYYDFSYSGFGEPPYISGSHGLGDLDDAVVSHRGEEDPYDEDTIRLSI 187
QY 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTRDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 240
DB 181 IGRPNVGSLLVAIIGEDRVIVSNVAGTTRDAIDTEYSYDGGDYVLIDTAGMRKKGKXY 247
QY 241 ESTEKYSVLRAKAIERSNVVLVDAEGLIEQDKRVAGYAHEOGKAVVIVNKKMDTYE 300
DB 241 ESTEKYSVLRAKAIERSNVVLVDAEGLIEQDKRVAGYAHEOGKAVVIVNKKMDTYE 307
QY 301 KOSKTKKKFEDEVRKKEFOFLDYAQIAFVSAKERTRLTLTPYINSESENHKKRVOSSTLN 360
DB 301 KOSKTKKKFEDEVRKKEFOFLDYAQIAFVSAKERTRLTLTPYINSESENHKKRVOSSTLN 367

QY 361 EVYDAISMNPPTDKGRRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 420
|||||
Db 368 EVYDAISMNPPTDKGRRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 427
QY 421 FGEFSTPIHIIARRK 436
|||||
Db 428 FGEFSTPIHIIARRK 443

RESULT 6

ABBA49212 ID ABBA49212 standard; Protein: 436 AA.

XX AC ABBA49212;

XX DT 05-FEB-2002 (first entry)

XX DE Listeria monocytogenes protein #1916.

XX KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

XX KM vitamin B12; bacterial infection; disease.

XX OS Listeria monocytogenes.

XX PN WO200177335-A2.

XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001WO-FR01118.

XX PR 11-APR-2000; 2000FR-0004629.

XX PA (INSP) INST PASTEUR.

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Eslihi H, Dehoux P,

PI Dusserget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P,

PI Danlens J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Madueno E, De Pablo B, Wehlund J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;

XX DR WPI; 2002-010914/01.

XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

XX PR and prevention of Listeria and related bacterial infections, and

XX PS related polypeptides

XX PS Claim 6; SEQ ID NO 1917; 192pp; French.

XX CC The present invention relates to the genome sequence of Listeria

XX CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of

XX CC it are useful for selecting probes and primers for detecting genes in L.

XX CC monocytogenes and related organisms, and for studying genetic

XX CC polymorphisms and other genomes. The present sequence is a protein

XX CC encoded from the genome sequence of the present invention. Proteins

XX CC expressed from the genome sequence are useful for raising specific

XX CC antibodies, identification of L. monocytogenes and related organisms, and

XX CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

XX CC B12. The genome sequence and proteins encoded by it are also useful for

XX CC selecting compounds that regulate gene expression and cell replication

XX CC and modulate L. monocytogenes-related diseases. In addition, the genome

XX CC sequence and proteins encoded by it are useful in pharmaceutical and

XX CC vaccines compositions for the treatment or prevention of infections by L.

XX CC monocytogenes and related organisms.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX CC Sequence 436 AA;

XX CC Query Match 78.2%; Score 1731; DB 23; Length 436;

Best Local Similarity 73.6%; Pred. No. 4,4e-142;
Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

QY 1 MTKPIVAIVGRPNVSKSTIFNRIIGERSYIEDTGVYRDRIRYSGSEMLTHDFNIIDGG 60
|||:|||||
Db 1 MAKPVVAIVGRPNVSKSTIFNRIIGERSYIEDTGVYRDRIRYSNEMGKRFNIIDGG 60
|||:|||||
QY 61 IEIGDAPFOTQIRAOAEIADIDADVITFMVNVREGLTQSDENVAQILVKKSPVLAANK 120
|||:|||||
Db 61 IDLSDEPFEQIRAOAEIADIDADVITFMVNVREGLTQSDENVAQILVKKSPVLAANK 120
|||:|||||
QY 121 VDNMEMRTDYYDYFSLGGEPEYPIGSHGLGLDLDVAVSHFGEEDDPYDEDITRLSI 180
|||:|||||
Db 121 VDNMEMRTDYYDYFSLGGEPEYPIGSHGLGLDLDVAVSHFGEEDDPYDEDITRLSI 180
|||:|||||
QY 181 IGRPNVGSLLVNAIILGEDRYVSNVAGTTRDAIDTEYSYDGDYVLIIDTAGMRKKGVY 240
|||:|||||
Db 181 IGRPNVGSLLVNAIILGEDRYVSNVAGTTRDAIDTEYSYDGDYVLIIDTAGMRKKGVY 240
|||:|||||
QY 241 ESTEYSVLRLAKAIERSNVVLYVIDAEGITIEDDKRVAGVAHEGKAVVIVNKMDFYE 300
|||:|||||
Db 241 ESTEYSVLRLAKAIERSNVVLYVIDAEGITIEDDKRVAGVAHEGKAVVIVNKMDFYE 300
|||:|||||
QY 301 KDSKTKMKFEDEVRKEFQFLDYAQIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
|||:|||||
Db 301 KDSKTKMKFEDEVRKEFQFLDYAQIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
|||:|||||
QY 361 EVYDAISMNPPTDKGRRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 420
|||||
Db 361 EVYDAISMNPPTDKGRRLNVFATQVAIEPPTFVFNVDVLMHFSYKRYLENOIRAA 420
|||||
QY 421 FGEFSTPIHIIARRK 435
|||||
Db 421 FGEFSTPIHIIARRK 435

RESULT 7

ABP27827 ID ABP27827 standard; Protein: 436 AA.

XX AC ABP27827;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polyptide SEQ ID NO 4830.

XX KW Streptococcus; GAS; GBs; group B streptococcus; Streptococcus agalactiae;

XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;

XX OS Streptococcus pyogenes.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PI (GENO-) INST GENOMIC RES.

XX PI Telford J, Masiqani V, Margalit Ros YI, Grandi G, Fraser C;

XX PT Tettelin H;

XX DR WPI; 2002-352536/38.

XX DR N-PSDB: ABN68458.

XX CC New Streptococcus protein for the treatment or prevention of infection

XX CC or disease caused by Streptococcus bacteria, such as meningitis, and

RESUR J 9
 ABP27826
 ID ABP27826 standard; Protein; 436 AA.
 XX
 AC ABP27826;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 4828.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Maignanl V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tetsteln H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN68457.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 Claim 1; Page 3645; 4525p; English.
 XX
 PS The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 Sequence 436 AA;
 XX

Query Match	68.8%	Score 1523;	DB 23;	Length 436;
Best Local Similarly	66.3%	Pred. No. 5.8e-124;		
Matches 289; Conservative	72;	Mismatches 73;	Indels 2;	Gaps 2

[illegible]

Dd	61	IDDVAPRNEQKKHQADIAMTADYVIFVYSGKEGTTDADEYVSRLLKYTKNRYVIAVKN	120
Qy	121	VDNMKEKRDYVDYFSLGEGEPYPISGSHGLGDLDAVVSHPGEEDDPYDEDTIRLSI	180
Dd	121	VDNPEMRNDIYDFYSLGGLDYPPLPSVSHGIGGDILDAIVENLPEEENE-NPDIIRFSL	179
Qy	181	IGRPVNGSSLYNALIGEDRYIVSVNAGTTRDAIDTEY-SYGDQYVLIDTAGMRKKGY	239
Dd	180	IGRPVNGSSLYNALIGEDRYIVASPAAGTTRDAIDTFNDSQOQETMTIDTAGMRKSGV	239
Qy	240	YESTEKYVLRALKRIERSNVLYVYDAOGIIEODKRVAGYAHEDGKAVYIVNNKDPY	239
Dd	240	YENTEKYSVMRMRADIRSDVYLMVYMAEGIREYDKRFLAGFHNHETGKGIIVNNKMDTI	239
Qy	300	EKDSKTMKKFEDEVEKREFOLDYAOIAFYSAKERTLRPLFYINASENNHKRVQSSSTL	359
Dd	300	EKDNTITVQWEMADIDYNQOFLSYAPIIFVSAETKQYLHKLPRMIKRISESQMKRIFSAVL	359
Qy	360	NEVYTDIAISMNETPDKGGRNLNVFYATOVAIEBPPTFVFEVNDVLELHFYSKYRLAQIRA	419
Dd	360	NDVIMDAIAINPTPDKGKRLKIFVATQVAVKPRFVEVNEBELMHFSYLRLENQIRE	419
Qy	420	AFGEFGTPIHIIARKR	435
Dd	420	AFVEFGTPIINLIARKR	435

RESULT 10
 AAB04108
 ID AAB04108 standard; Protein; 436 AA.
 XX
 AC AAB04108;
 XX
 DT 11-APR-2001 (first entry)
 DE yphc protein of Streptococcus pneumoniae (GTP binding protein).
 XX
 KW yphc; GTP binding protein; antibody; treatment; infection;
 KW bacteraemia; otitis media; conjunctivitis; pneumoniae;
 XX meningitis; sinusitis; pleural empyema; endocarditis.
 OS Streptococcus pneumoniae.
 PN WO200068427-A1.
 PD 16-NOV-2000.
 PF 02-MAY-2000; 2000MO-US11894.
 PR 07-MAY-1999; 99US-0307003.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Zalcasin M, Biswas S, Burnham MKR, Warren PV, Ingram KA;
 PI Chaliker AF, So CT, Holmes DJ, Warren RL, Van Horn S, Traini CM.
 DR WPI: 2000-687653/67.
 DR N-PSDB: AAA54516.
 PT Streptococcus pneumoniae yphc protein and DNA sequence, useful for
 PS treating infections, meningitis, and bacteraemia
 Claim 1: Page 37-38; 39pp: English.

A new nucleotide is described which encodes a 436 residue Streptococcus pneumoniae yphC (GTP binding protein). The DNA sequence can be used to transform a host cell to produce the protein and the products can be used to treat bacterial infections (especially streptococcus pneumoniae infections, and Helicobacter pylori infections), otitis media, conjunctivitis, pneumonitis, bacteremia, meningitis, sinusitis, pleural empyema, endocarditis. The encoded protein and its peptide fragments can be

CC used to generate antibodies.

XX Sequence 436 AA;

Query Match 68.8%; Score 1522; DB 21; Length 436;
Best Local Similarity 66.2%; Pred. No. 7,1e-124;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

QY 1 MRPPIAIVRPNVSKSTINRIVGEVSVIEDTPGVTDRIRIYSGGWLTHDNIIDTGG 60
1 MRPPIAIVRPNVSKSTINRIVGEVSVIEDTPGVTDRIRIYSGGWLTHDNIIDTGG 60
DB 1 MALPTAIIVRPNVSKSTINRIVGEVSVIEDTPGVTDRIRIYSGGWLTHDNIIDTGG 60
QY 61 IEIGDAPFQOIRAOEIADEADVIFMVNVEGLTSDENVAQILYKSKRPVLAIVNK 120
61 IDVDAPFMEQIKHQAIEAEADVIFVVSKEGIDDAEYARUKYKHKRVYLAIVNK 120
QY 121 VDNEMKRTDYDYFSLGFGEPYRISGSHGLDLDVAVSHGE--EEDPYDEDTIRL 178
121 VDNEMKRTDYDYFSLGFGEPYRISGSHGLDLDVAVSHGE--EEDPYDEDTIRL 178
DB 121 VDNEMKRTDYDYFSLGFGEPYRISGSHGLDLDVAVSHGE--EEDPYDEDTIRL 178
QY 179 SIIGRPVNGKSSLVNAILGDRIVSNVAGTTRDAIDTEVS-YDGQDYVLIDTAGMKKGG 237
179 SIIGRPVNGKSSLVNAILGDRIVSNVAGTTRDAIDTEVS-YDGQDYVLIDTAGMKKGG 237
DB 178 SLIGRPVNGKSSLVNAILGDRIVSNVAGTTRDAIDTEVS-YDGQDYVLIDTAGMKKGG 237
QY 238 KYESTEKYSVLRAKAIERSNVLYVDAEOGIIEDOKRVAGYAHOGKAVIYVNVKM 297
238 KYESTEKYSVLRAKAIERSNVLYVDAEOGIIEDOKRVAGYAHOGKAVIYVNVKM 297
DB 238 KYESTEKYSVLRAKAIERSNVLYVDAEOGIIEDOKRVAGYAHOGKAVIYVNVKM 297
QY 298 TVEKDSKTMKKFEDEVAKREFQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVQSS 357
298 TVEKDSKTMKKFEDEVAKREFQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVQSS 357
DB 298 TVEKDSKTMKKFEDEVAKREFQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVQSS 357
QY 358 TLNEVYTDASNPPTDKGRRLNVFATQVALEPPFVAVVNDVLMHSYKRYLENOI 417
358 TLNEVYTDASNPPTDKGRRLNVFATQVALEPPFVAVVNDVLMHSYKRYLENOI 417
DB 358 TLNEVYTDASNPPTDKGRRLNVFATQVALEPPFVAVVNDVLMHSYKRYLENOI 417
QY 418 RAAFGEGTPPIHIIARKR 435
418 RAAFGEGTPPIHIIARKR 435
DB 418 RAAFGEGTPPIHIIARKR 435

RESULT 11
AA81730
ID AAY81730 standard; Protein: 436 AA.

XX AAY81730:

02-JUN-2000 (first entry)

Streptococcus pneumoniae protein sequence ID36.

XX Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
KW pneumococcal septicaemia; sinusitis; meningitis; therapy.

OS Streptococcus pneumoniae.

XX WO200006738-A2.

PD 10-FEB-2000.

PF 27-JUL-1999; 99MO-GB02452.

PR 27-JUL-1998; 98GB-0016336.

PR 19-MAR-1999; 99US-0125329.

XX (MICR-) MICROBIAL TECHNIQS LTD.

PI Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;

XX WPI; 2000-195301/17.

DR N-PSDB; AA291826.

XX Streptococcal proteins and polynucleotides useful for diagnosis,
PT treatment and prophylaxis of bacterial infections
XX
PS Claim 2; Page 49; 76pp: English.

CC This sequence represents a Streptococcus pneumoniae protein of the
CC invention. The proteins (or their homologues, derivatives and/or
CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
CC compositions comprising the proteins are useful as vaccines and also in
CC diagnostic assays. The sequences are useful for the detection or
CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
CC with them. Agents capable of antagonising, inhibiting or interfering with
CC the function or expression of the protein or polypeptide are useful in
CC medical compositions in the treatment or prophylaxis of S. pneumoniae
CC infection. As the sequences can be used to treat S. pneumoniae infection,
CC they can be used to treat bacterial pneumonia, which has high rates in
CC young children, the elderly, and in patients with predisposing conditions
CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
CC or with immunosuppressive disorders, especially AIDS. They can also be
CC used to treat pneumococcal septicaemia, otitis media, sinusitis, and
CC meningitis.

SO Sequence 436 AA;

Query Match 68.8%; Score 1522; DB 21; Length 436;
Best Local Similarity 66.2%; Pred. No. 7,1e-124;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

QY 1 MRPPIAIVRPNVSKSTINRIVGEVSVIEDTPGVTDRIRIYSGGWLTHDNIIDTGG 60
1 MRPPIAIVRPNVSKSTINRIVGEVSVIEDTPGVTDRIRIYSGGWLTHDNIIDTGG 60
DB 1 MALPTAIIVRPNVSKSTINRIVGEVSVIEDTPGVTDRIRIYSGGWLTHDNIIDTGG 60
QY 61 IEIGDAPFQOIRAOEIADEADVIFMVNVEGLTSDENVAQILYKSKRPVLAIVNK 120
61 IDVDAPFMEQIKHQAIEAEADVIFVVSKEGIDDAEYARUKYKHKRVYLAIVNK 120
DB 61 IDVDAPFMEQIKHQAIEAEADVIFVVSKEGIDDAEYARUKYKHKRVYLAIVNK 120
QY 121 VDNEMKRTDYDYFSLGFGEPYRISGSHGLDLDVAVSHGE--EEDPYDEDTIRL 178
121 VDNEMKRTDYDYFSLGFGEPYRISGSHGLDLDVAVSHGE--EEDPYDEDTIRL 178
DB 121 VDNEMKRTDYDYFSLGFGEPYRISGSHGLDLDVAVSHGE--EEDPYDEDTIRL 178
QY 179 SIIGRPVNGKSSLVNAILGDRIVSNVAGTTRDAIDTEVS-YDGQDYVLIDTAGMKKGG 237
179 SIIGRPVNGKSSLVNAILGDRIVSNVAGTTRDAIDTEVS-YDGQDYVLIDTAGMKKGG 237
DB 178 SLIGRPVNGKSSLVNAILGDRIVSNVAGTTRDAIDTEVS-YDGQDYVLIDTAGMKKGG 237
QY 238 KYESTEKYSVLRAKAIERSNVLYVDAEOGIIEDOKRVAGYAHOGKAVIYVNVKM 297
238 KYESTEKYSVLRAKAIERSNVLYVDAEOGIIEDOKRVAGYAHOGKAVIYVNVKM 297
DB 238 KYESTEKYSVLRAKAIERSNVLYVDAEOGIIEDOKRVAGYAHOGKAVIYVNVKM 297
QY 298 TVEKDSKTMKKFEDEVAKREFQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVQSS 357
298 TVEKDSKTMKKFEDEVAKREFQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVQSS 357
DB 298 TVEKDSKTMKKFEDEVAKREFQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVQSS 357
QY 358 TLNEVYTDASNPPTDKGRRLNVFATQVALEPPFVAVVNDVLMHSYKRYLENOI 417
358 TLNEVYTDASNPPTDKGRRLNVFATQVALEPPFVAVVNDVLMHSYKRYLENOI 417
DB 358 TLNEVYTDASNPPTDKGRRLNVFATQVALEPPFVAVVNDVLMHSYKRYLENOI 417
QY 418 RAAFGEGTPPIHIIARKR 435
418 RAAFGEGTPPIHIIARKR 435
DB 418 RAAFGEGTPPIHIIARKR 435

RESULT 12
ABU02185
ID ABU02185 standard; Protein: 436 AA.

XX ABU02185:

XX 11-FEB-2003 (first entry)

DE S. pneumoniae type 4 strain protein from coding region #1762.

XX Bacterial meningitis; pneumonia; sepsis; otitis media;

XX	ear infection; antiinflammatory; antibacterial; immunostimulant;
KW	auditory; respiratory; gene therapy; vaccine.
XX	
OS	Streptococcus pneumoniae type 4 strain.
XX	
PN	WO200277021-A2.
XX	
PD	03-OCT-2002.
XX	
PF	27-MAR-2002; 2002WO-IB02163.
XX	
PR	27-MAR-2001; 2001GB-0007658.
XX	
PA	(CHIR-) CHIRON SPA.
XX	(GENO-) INST GENOMIC RES.
PI	
XX	Maignani V, Tettelin H, Fraser C;
DR	MP1: 2003-040579/03.
XX	
DR	N-PSDB: ABX07474.
XX	
PT	New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT	useful as medicaments for treating or preventing a disease or infection
PT	due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
PT	or ear infection -
XX	
XX	Claim 1; SEQ ID NO 3524; 56pp; English.

[illegible][illegible]

XX	RESULT 13
XX	ABB54068
XX	ID ABB54068 standard; Protein; 436 AA.
XX	AC ABB54068;
XX	DT 16-MAY-2002 (first entry)
XX	DE Lactococcus lactis protein yphL.
KM	Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX	XX
XX	XX Lactococcus lactis IL1403.
XX	XX
XX	XX FR2807446-A1.
PD	12-OCT-2001.
XX	XX
XX	XX 11-APR-2000; 2000FR-0004630.
XX	XX
XX	XX 11-APR-2000; 2000FR-0004630.
XX	XX
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.
XX	XX
PI	Bolotline A, Sorokline A, Renault P, Ehrlich SD;
XX	XX
DR	WPI; 2002-043418/06.
XX	XX
PT	New nucleotide sequence useful in the identification or Lactococcus
PT	lactis and related species -
XX	XX
PS	Claim 6; SEQ ID No 770; 2504pp; French.
XX	XX
CC	The present invention is related to a Lactococcus lactis nucleotide
CC	sequence (AB90521) and related proteins (AB853300-AB855621). The
CC	nucleic acid sequence is useful in the detection and/or amplification of
CC	nucleic acid sequence, particularly to identify Lactococcus lactis or
CC	related species. The proteins of the invention are useful for the
CC	biosynthesis or biodegradation of a composition of interest. The
CC	invention helps research in lactic bacteria, particularly useful in the
CC	production of yogurt and cheese.
CC	Note: The sequence data for this patent is based on equivalent patent
CC	WO200177314 (published 18-OCT-2001) which is available in electronic
CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	XX
XX	Sequence 436 AA;

Query Match 68.7%; Score 1520; DB 23; Length 436;
Best Local Similarity 66.3%; Pred. No. 1.1e-123;
Matches 289; Conservative 69; Mismatches 76; Indels 2; Gaps 2;

QY 1 MTKPIVALVGRPNVCKSTIFNRIVGERVSIYEDTFCVTRDRIRYSSGEMLTHTDFNITIDGG 60
DB 1 MSLPIVALVGRPNVCKSTIFNRIVGERVSIYEDTFCVTRDRIRYATGEMLTHTKFNITIDGG 60
QY 61 IEIGDAPFQTOIRAQAEIAIDADVIFMVNVRGLTQSDENVAOILYKSKRPVLAVNK 120
DB 61 IELSDPEPMTFIRAQAEIATMEADVIIVAVDGETGITDADEVANILYRTOKPVLIVVVK 120
QY 121 VDNMFMRTDVFYSLGCEPPISSGSHGLGLDLDLVAVSHFGEEDDPDEDITRLSI 180
DB 121 VDNPMRMEIFDFYSLGLDPPYSAVHIGTGDLDAIVONLPNEIEEE-NEVYIKFSL 179
QY 181 IGRPNVGSLLVNAITLGEDRVIVSNAGTRDAIDTEX-SYDGODVYLDIPAGMKKCKV 239
DB 180 IGRPNVGSLLVNAITLGEDRVIVSNAGTRDAIDTEX-SYDGODVYLDIPAGMKKCKV 239
QY 240 YESTERKYSVLAKAIEKRSNVLVVIADEQIIEODKRVAGYAHBOGRAVIVVKNKMDTV 299
DB 240 YENFEKYSVNRAMRAIDRSIDIVLVYNAEGRIDRMIRAGFAHAGGILIVVKNKMDTV 299
QY 300 EKDSITMKKFEDEVKKEQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVOSSTL 359
DB 300 EKDNVTMNFLEITRTKRFIDYAPIVYVSAKTGRLKIKRHHAAONLRISSSVL 359
QY 360 NEVVYDAISMNPTPTDKGRNLNVFATQVAIEPPEVFNVDVLMHFSYKRYLENOIRA 419
DB 360 NDVIMDAVINPTPTDKGRNLKIFATQVAIKPPEVFNVDVLMHFSYKRYLENOIRK 419
QY 420 AFEGEGTPIHIIARKR 435
DB 420 AFEGEGTPIHIIARKR 435

RESULT 14
AAG99939 ID AAG99939 standard; Protein: 448 AA.
XX AAG99939;
XX 27-SEP-2001 (first entry)
XX
XX ERA binding domain polypeptide SEQ ID NO 381.
XX
XX ERA binding domain; Escherichia coli; GTPase; antimicrobial;
KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
XX peptide therapy.
XX
XX Mycobacterium leprae.
XX
XX WO200153458-A2.
XX
XX 26-JUL-2001.
XX
XX 17-JAN-2001; 2001WO-US01786.
XX
XX 18-JAN-2000; 2000US-0176870.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Lupas AN, Pearce KH;
XX
XX WPI: 2001-476108/51.
XX
XX New ERA binding domain polypeptides and polynucleotides encoding them,
XX useful as research reagents and materials for discovery of treatments
XX and diagnostics for diseases, or for genetic immunisation -
XX
XX Claim 1; Page 52-53; 27pp; English.

XX The present invention relates to ERA binding domain polypeptides
CC (AAG99559-AAG99989 and AAM00010-AAM00021). The era gene in Escherichia
CC coli codes for an essential GTPase protein able to autophosphorylate at
CC serine and/or threonine residues. The protein has potential antimicrobial
CC and antibacterial activity and is useful in screening for antagonists,
CC agonists and for compounds with antibiotic activity. The proteins are
CC also useful in determining their role in pathogenesis of infection,
CC dysfunction and disease and could be used as part of a vaccine and/or
CC peptide therapy.
SQ Sequence 448 AA:

Query Match 44.9%; Score 933; DB 22; Length 448;
Best Local Similarity 45.6%; Pred. No. 8.7e-78;
Matches 199; Conservative 93; Mismatches 136; Indels 8; Gaps 6;

QY 1 MTKPIVALVGRPNVCKSTIFNRIVGERVSIYEDTFCVTRDRIRYSSGEMLTHTDFNITIDGG 60
DB 1 MSLPIVALVGRPNVCKSTIFNRIVGERVSIYEDTFCVTRDRIRYATGEMLTHTKFNITIDGG 58
QY 61 IEIG-DAPFQTOIRAQAEIAIDADVIFMVNVRGLTQSDENVAOILYKSKRPVLAVN 119
DB 59 LVFNDDSEFLPEIRQANLA--EAKAALFVVDGQGPAPASDEITAOILRQOSVPIVLAVN 116
QY 120 KVDNMEM-RTDVFYSLGCEPPISSGSHGLGLDLDLVAVSHFGEEDDPDEDITRL 178
DB 117 KCESPDGAIQAAEFMHGLGEPYMSAIGHSGTGLDLDLLEYLPAPQEP-EEDEIKV 175
QY 179 SIIGRPNVGKSSLVNAITLGEDRVIVSNAGTRDAIDTEXSYDGODVYLDITAGARKKCK 238
DB 176 AIVGRPNVGSLLVNAITLGEDRVIVSNAGTRDAIDTEXSYDGODVYLDITAGARKKCK 235
QY 239 YESTERKYSVLAKAIEKRSNVLVVIADEQIIEODKRVAGYAHBOGRAVIVVKNKMDT 298
DB 236 VDYGAEFGINRAFAIRADVYLFVLDVLDGCVTQDLKLAGRIIEDRAVAVVINKMDA 295
QY 299 VEKDSITMKKFEDEVKKEQFLDYAQAIAFVSAKERTRLTLPYINEASENHKKRVOSST 358
DB 296 VEKDSYTIYEHREQLMARLYFMDVMAEMIFVSAQTGLRVQKILDCVDIAAOCHRRRVYTA 355
QY 359 LNEVVYDAISMNPTPTDK-GRRLNVFATQVAIEPPEVFNVDVLMHFSYKRYLENOI 417
DB 356 INEVLEAVSWHSPPTTRQKQKGIYGVSTOPOPAIALFVNDPNRPNRYRIEKO 415
QY 418 RAAREGECTPIHIIAR 433
DB 416 RKQLGFGSPTRLFWR 431

RESULT 15
AAG99941 ID AAG99941 standard; Protein: 448 AA.
XX AAG99941;
XX 27-SEP-2001 (first entry)
XX
XX ERA binding domain polypeptide SEQ ID NO 383.
XX
XX ERA binding domain; Escherichia coli; GTPase; antimicrobial;
KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
XX peptide therapy.
XX
XX Treponema pallidum.
XX
XX WO200153458-A2.
XX
XX 26-JUL-2001.
XX
XX 17-JAN-2001; 2001WO-US01786.
XX
XX 18-JAN-2000; 2000US-0176870.

XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Lupas AN, Pearce KH;
XX WPI: 2001-476108/51.
DR
XX
XX New ERA binding domain polypeptides and polynucleotides encoding them,
PT useful as research reagents and materials for discovery of treatments
PT and diagnostics for diseases, or for genetic immunisation -
XX
PS Claim 1; Page 53; 279pp: English.
XX
XX The present invention relates to ERA binding domain polypeptides
CC (AAG9559-AAG9989 and AAM00010-AAM00021). The era gene in Escherichia
CC coli codes for an essential GTPase protein able to autophosphorylate at
CC serine and/or threonine residues. The protein has potential antimicrobial
CC and antibacterial activity and is useful in screening for antagonists,
CC agonists and for compounds with antibiotic activity. The proteins are
CC also useful in determining their role in pathogenesis of infection.
CC dysfunction and disease and could be used as part of a vaccine and/or
CC peptide therapy.
XX
XX Sequence 448 AA;
SQ
Query Match: 42.3%; Score 935.5; DB 22; Length 448;
Best Local Similarity 43.2%; Pred. No. 8.8e-73;
Matches 189; Conservative 86; Mismatches 154; Indels 9; Gaps 4;
QY 6 VAIIVRPNVGKSTIFNRIYGERVSIYEDTPGVTDRDIYSSGEMLTDFNIIDTGGIEIGD 65
DB 4 VAIIGRTNVGKSTIFNRIYGERVSIYEDTPGVTDRDIYSSGEMLTDFNIIDTGGIEIGD 63
QY 66 APEGTOIRAOAEIIDEADVIIFMWNVREGITQSDENVAQILYKSK-KPVVLAIVKVDNM 124
DB 64 TPLOQLALQVQALISQAKIIFLVSLQEQIUNSDDFYAKVLKKNKDKPILVYNKAENF 123
QY 125 EMRT---DYVDYFSLGFGEPYPISGSHGLGDLIDAVVSH---FGEEEDPYDEDTRL 178
DB 124 NPKTAETLNDYISLGFGRPVYISAHIGIGIDMLDLVKONOLLPNENND--DIAKITRF 181
QY 179 SIIGRPNVGKSSILVNAIIGEDRIVSNVAGTTTDAIDTEYSYDGDVYLIDTAGMKRKGK 238
DB 182 CVIGKPNVGKSSILNQVYKQKRVLVNSGTTTDAIDVPLKVNCEKFLIDTAGIKRKGK 241
QY 239 VYESTEKYSVLAKAIEKRSNVVLVIDAEGITIEODKRVAGYAHGCKAVIYVKNKWD 298
DB 242 INMGIEIYASYIKTILAIRSNVILIMWDGSKPISEODEVIGLAQALIPVILVKNKWD 301
QY 299 VEKDSKTMKKRFEDEVKKEFOFLDYAQIAFVSAKERTRLTLPYINEASENHKKRVQSS 358
DB 302 VLKNNNTNMYKKMLKLFKHLDPALFISVLKNQRLNTLFEOLKITQSOLETVATPL 361
QY 359 LNEVVTDAISMNPPTDKGRRLNLFYATQVAIEPTEVFVNDVELMHFSYKRYLENQIR 418
DB 362 LNDVYQQAQLNQPPLFKGRRLQITTYAVQTSQIPIHFVLPFCNDPKYILHFSYARLENKIR 421
QY 419 AAFGEETPIHITAKRN 436
DB 422 ENFGFNSVPIISLYFKSKN 439

Search completed: September 10, 2003, 00:32:44
Job time : 88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:29:51 ; Search time 43 Seconds

(without alignments)
975.106 Million cell updates/sec

Title: US-09-815-242-12600

Perfect score: 2213

Sequence: 1 MTKPIVAIVGRPNVGRKSTIF.....IRAFGEFGPIHIIAKRN 436

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2213	100.0	436	2 C89926	hypothetical prote
2	1731	78.2	436	2 A11316	hypothetical prote
3	1724	77.9	436	2 A11688	hypothetical prote
4	1699	76.8	436	2 A69936	probable GTP bindi
5	1658	74.9	437	2 F83854	hypothetical prote
6	1522	68.8	436	2 B95199	phosphoglycerate d
7	1522	68.8	436	2 H98065	phosphoglycerate d
8	1520	68.7	436	2 C86719	GTP-binding protei
9	1218.5	55.1	438	2 B97111	probable GTPase (I
10	1037	46.9	453	2 AB1867	GTP binding protei
11	1021	46.1	442	2 S75645	probable GTP bindi
12	995	45.0	442	2 B82899	conserved GTP bindi
13	984.5	44.5	435	2 H90522	gtp-binding protei
14	935.5	42.3	448	2 D64236	hypothetical prote
15	923	41.7	503	2 E91050	probable GTP-bind
16	923	41.7	503	2 A85895	probable GTP-bind
17	923	41.7	503	2 F65027	probable GTPase/GT
18	910	41.1	490	2 AF0821	GTP-binding protei
19	906.5	41.0	490	2 B82284	GTP-binding protei
20	896.5	40.5	495	2 AC0350	probable GTP-bind
21	864.5	39.1	433	2 A70465	essential GTP bindi
22	860.5	38.9	485	2 G81149	probable GTPase N
23	858.5	38.8	485	2 E81871	probable GTP-bind
24	852	38.5	504	2 F64143	probable GTPase/GT
25	842.5	38.1	433	2 B72253	conserved GTP-bind
26	840.5	38.0	493	2 B83171	GTP-binding protei
27	838	37.9	493	2 D87454	GTP-binding protei
28	806.5	36.4	463	2 H70504	probable GTP-bind
29	797.5	36.0	483	2 AH3445	GTP-binding protei

30	789	35.7	476	2 A97636	probable GTP-bind
31	789	35.7	476	2 AC2859	GTP-binding protei
32	784	35.4	438	2 F75290	conserved GTP-bind
33	782.5	35.4	461	2 S72953	probable GTP-bind
34	782.5	35.4	462	2 F87080	probable GTP-bind
35	779	35.2	465	2 D82804	GTP-binding protei
36	778.5	35.2	476	2 A95982	probable GTP-bind
37	752	34.0	453	2 C85000	hypothetical GTP-b
38	746.5	33.7	447	2 F97828	conserved GTP-bind
39	737.5	33.3	447	2 H71673	hypothetical prote
40	734.5	33.2	460	2 H81381	probable GTP-bind
41	719.5	32.5	490	2 A81744	GTP-binding protei
42	710.5	32.1	490	2 H71480	probable GTPase/GT
43	706	31.9	458	2 B64624	GTP-binding protei
44	701	31.7	462	2 E71891	probable GTP bindi
45	685.5	31.0	487	2 B86596	GTPase/GTP-binding

ALIGNMENTS

RESULT 1									
C89926									
hypothetical protein SA1307 [imported] - Staphylococcus aureus (strain N315)									
C:Species: Staphylococcus aureus									
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001									
C:Accession: C89926									
R:Kuroda, M.; Ohta, T.; Uchlyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cul, L.; O									
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.									
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsutsu, K.									
Lancet 357, 1225-1240, 2001									
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.									
A:Reference number: A89758; MUID:21311952; PMID:11418146									
A:Accession: C89926									
A:Status: Preliminary									
A:Molecule type: DNA									
A:Residue: 1-436 <KUR>									
A:Cross-references: GB:BA000018; PID:q13701274; PIDN:BAH42568.1; GSPDB:GN00149									
A:Experimental source: strain N315									
A:Gene: SA1307									
C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat									
Query Match									
Best Local Similarity 100.0%; Score 2213; DB 2; Length 436;									
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
OY	1	MTKPIVAIVGRPNVGRKSTIFNRIYGERYSIVEDPFGVTRDRIRYSSGEMLTHDFNIIDTGC	60						
DB	1	MTKPIVAIVGRPNVGRKSTIFNRIYGERYSIVEDPFGVTRDRIRYSSGEMLTHDFNIIDTGC	60						
OY	61	IEIGDAPFOQIIRAQAEIAIDEDAVIIFMVNVRBSLTQSDENVAOILYKSKPPVLAANK	120						
DB	61	IEIGDAPFOQIIRAQAEIAIDEDAVIIFMVNVRBSLTQSDENVAOILYKSKPPVLAANK	120						
OY	121	VDNMEMRDVDFYSLGEGEPYISGSLGLGDLDAVVSFGSEEDPEYEDTRLSI	180						
DB	121	VDNMEMRDVDFYSLGEGEPYISGSLGLGDLDAVVSFGSEEDPEYEDTRLSI	180						
OY	121	VDNMEMRDVDFYSLGEGEPYISGSLGLGDLDAVVSFGSEEDPEYEDTRLSI	180						
DB	121	VDNMEMRDVDFYSLGEGEPYISGSLGLGDLDAVVSFGSEEDPEYEDTRLSI	180						
OY	181	IGRPVNGKSSLVNALLGEDRVIVSNVAGTTRDAIDTEXSYSDQDVLIDTAGMRKKGVY	240						
DB	181	IGRPVNGKSSLVNALLGEDRVIVSNVAGTTRDAIDTEXSYSDQDVLIDTAGMRKKGVY	240						
OY	241	ESTEKYSVLRALKAIERSNVVLVDAEAGIIEQDKRVAAGYAHDEGKAVIYVNWDTVE	300						
DB	241	ESTEKYSVLRALKAIERSNVVLVDAEAGIIEQDKRVAAGYAHDEGKAVIYVNWDTVE	300						
OY	301	KDSKTMKKFEDEVREKFOFLDVAQIAFASAKERTLRLLPYINSESHKKRVOSTLN	360						
DB	301	KDSKTMKKFEDEVREKFOFLDVAQIAFASAKERTLRLLPYINSESHKKRVOSTLN	360						
OY	361	EVVTAISMNPTPTKGRRLNVFATQVAIEBPFTVVVNDVLMHFSKRYLENOIRAA	420						
DB	361	EVVTAISMNPTPTKGRRLNVFATQVAIEBPFTVVVNDVLMHFSKRYLENOIRAA	420						

Db 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 Oy 421 FGEGTPIHIARKN 436
 Db 421 FGEGTPIHIARKN 436

RESULT 2

hypoethetical protein lmo1937 [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: A11316
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; M0ID:21537279; PMID:11679669
 A:Accession: A11316
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00015.1; PID:g16411390; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1937
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation

Query Match 78.28; Score 1731; DB 2; Length 436;
 Best Local Similarity 73.68; Pred. No. 4,3e-99;
 Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

Oy 1 MTKPIVAIGRPVNGKSTIFNRIVGRVSVIEDPTGVTRDRISSEGLTHDNIITDGG 60
 1 MAKPVVAIGRPVNGKSTIFNRIVGRVSVIEDPTGVTRDRISSEGLTHDNIITDGG 60
 Db 1 MAKPVVAIGRPVNGKSTIFNRIVGRVSVIEDPTGVTRDRISSEGLTHDNIITDGG 60
 Oy 61 IEIGDAPFOTQIRAOAEIAIDEDAVIIFMNVNREGTOSDEMAVQILYKSKRPVLAIVK 120
 1 IEIGDAPFOTQIRAOAEIAIDEDAVIIFMNVNREGTOSDEMAVQILYKSKRPVLAIVK 120
 Db 61 IEIGDAPFOTQIRAOAEIAIDEDAVIIFMNVNREGTOSDEMAVQILYKSKRPVLAIVK 120
 Oy 121 VNMENRTVDYFYSLSGFEPPYPISSGSHGLGDLDAVSHFGEEDPDYEDTIRLSI 180
 121 VNMENRTVDYFYSLSGFEPPYPISSGSHGLGDLDAVSHFGEEDPDYEDTIRLSI 180
 Db 121 VNMENRTVDYFYSLSGFEPPYPISSGSHGLGDLDAVSHFGEEDPDYEDTIRLSI 180
 Oy 181 IGRPNVNGKSSILNALGEDRVIVSNVAGTTRDAIDTEYSYDGDVYLIDTAGMRKRGKY 240
 181 IGRPNVNGKSSILNALGEDRVIVSNVAGTTRDAIDTEYSYDGDVYLIDTAGMRKRGKY 240
 Db 181 IGRPNVNGKSSILNALGEDRVIVSNVAGTTRDAIDTEYSYDGDVYLIDTAGMRKRGKY 240
 Oy 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHEOGKAVVIVNKKDVE 300
 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHEOGKAVVIVNKKDVE 300
 Db 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHEOGKAVVIVNKKDVE 300
 Oy 301 KDSKTKKREDEVRKPEFOLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
 301 KDSKTKKREDEVRKPEFOLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
 Db 301 KDSKTKKREDEVRKPEFOLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
 Oy 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 Db 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 Oy 421 FGEGTPIHIARKN 435
 421 FGEGTPIHIARKN 435
 Db 421 FGEGTPIHIARKN 435

RESULT 3

hypoethetical protein homolog lln2051 [imported] - Listeria innocua (strain Clp11262)
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

data no good

C:Accession: A11688
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; M0ID:21537279; PMID:11679669
 A:Accession: A11688
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97281.1; PID:g16414552; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: lln2051
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat

Query Match 77.98; Score 1724; DB 2; Length 436;
 Best Local Similarity 73.18; Pred. No. 1,2e-98;
 Matches 318; Conservative 67; Mismatches 50; Indels 0; Gaps 0;

Oy 1 MTKPIVAIGRPVNGKSTIFNRIVGRVSVIEDPTGVTRDRISSEGLTHDNIITDGG 60
 1 MAKPVVAIGRPVNGKSTIFNRIVGRVSVIEDPTGVTRDRISSEGLTHDNIITDGG 60
 Db 1 MAKPVVAIGRPVNGKSTIFNRIVGRVSVIEDPTGVTRDRISSEGLTHDNIITDGG 60
 Oy 61 IEIGDAPFOTQIRAOAEIAIDEDAVIIFMNVNREGTOSDEMAVQILYKSKRPVLAIVK 120
 1 IEIGDAPFOTQIRAOAEIAIDEDAVIIFMNVNREGTOSDEMAVQILYKSKRPVLAIVK 120
 Db 61 IEIGDAPFOTQIRAOAEIAIDEDAVIIFMNVNREGTOSDEMAVQILYKSKRPVLAIVK 120
 Oy 121 VNMENRTVDYFYSLSGFEPPYPISSGSHGLGDLDAVSHFGEEDPDYEDTIRLSI 180
 121 VNMENRTVDYFYSLSGFEPPYPISSGSHGLGDLDAVSHFGEEDPDYEDTIRLSI 180
 Db 121 VNMENRTVDYFYSLSGFEPPYPISSGSHGLGDLDAVSHFGEEDPDYEDTIRLSI 180
 Oy 181 IGRPNVNGKSSILNALGEDRVIVSNVAGTTRDAIDTEYSYDGDVYLIDTAGMRKRGKY 240
 181 IGRPNVNGKSSILNALGEDRVIVSNVAGTTRDAIDTEYSYDGDVYLIDTAGMRKRGKY 240
 Db 181 IGRPNVNGKSSILNALGEDRVIVSNVAGTTRDAIDTEYSYDGDVYLIDTAGMRKRGKY 240
 Oy 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHEOGKAVVIVNKKDVE 300
 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHEOGKAVVIVNKKDVE 300
 Db 241 ESTEKYSVLRALKAIERSNVVLVIDAEOGITEODKRVAGVAHEOGKAVVIVNKKDVE 300
 Oy 301 KDSKTKKREDEVRKPEFOLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
 301 KDSKTKKREDEVRKPEFOLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
 Db 301 KDSKTKKREDEVRKPEFOLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTLN 360
 Oy 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 Db 361 EYVTDALISNNPTPTDGRRLNPFYATQVAIEPPTFVFVNDVDELHMFSTKRYLENDQIRAA 420
 Oy 421 FGEGTPIHIARKN 435
 421 FGEGTPIHIARKN 435
 Db 421 FGEGTPIHIARKN 435

RESULT 4

probable GTP binding protein yphc - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
 C:Accession: A69936
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emerson, P.T.; Entlian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galitz, A.; Gal lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scani A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

A:keuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, P.; Togononi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M.
A:Authors: Yoshikawa, H.F.: Zunstein, E.; Yoshikawa, H.; Danchin, A.
A:title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MWID:98044033; PMID:9384377
A:Accession: A69936
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-436 <KUN>
A:Cross-references: GB:J299115; GB:AL009126; NID:g2634478; PIDN:CAB14200.1; PID:g2634702
A:Experimental source: strain 168
C:Genetics:
A:Gene: yphc
C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
C:Keywords: duplication; GTP binding; nucleotide binding; P-loop
F:4-122/Domain: translation elongation factor Tu homology <ET1>
F:10-17/Region: nucleotide-binding motif A (P-loop)
F:119-122/Region: GTP-binding NKXD motif
F:145-147/Region: GTP-binding SAK/L motif #status atypical
F:176-297/Domain: translation elongation factor Tu homology <ET2>
F:182-189/Region: nucleotide-binding motif A (P-loop)
F:294-297/Region: GTP-binding NKXD motif
F:329-331/Region: GTP-binding SAK/L motif

Query Match	76.8%	Score 1699	DB 2	length 436
Best Local Similarity	73.6%	Pred. No. 4e-97		
Matches 320	Conservative 57	Mismatches 58	Indels 0	Gaps 0
QY	1	MTKPIVAIVGRPNVNGKSTINRIVGERVSIYEDTPPGVTRDRIRYSSGEMLTHDNLITDGG	60	
Db	1	MKRPVAIVGRPNVNGKSTINRIVGERVSIYEDTPPGVTRDRIRYSSAEMNLNDITDGG	60	
QY	61	IEIGDAPFQQTQIAQAQAIADQADVITFMVNVREGILQSDQENVAAQILYKSKRPVLAIVK	120	
Db	61	IDIGDEPFLAQIQQAQAIADQADVITFMVNVREGVTADEVAALILKTRKRPVLAIVK	120	
QY	121	VDNMERITDVYDFYSLGFGEPYPISSGSHGLGLDLDVAVSHFGGEEEDPPYEDYTRLSI	180	
Db	121	LNTETKRANIVDYFSLGFGEPYPISTGHGLGLDLDVAVAHEFKNIPEKYNVEVYQFLD	180	
QY	181	IGRPNVGKSSIVAAIGEDRVIVSNVAGTPTDADITDESYQGDVYLDITAGMRKKGKY	240	
Db	181	IGRPNVGKSSIVAAIGEEERVIVSNVAGTTRDADVTSFYVQQQEVYIDVTAGMRKKGY	240	
QY	241	ESTEKYSVLRALKAIERSNVVLVVIDAEQGLIEDDKRAGYAHBQGAIVIVNKKMDVE	300	
Db	241	ETTEKYSVLRALKAIDRESEVAVYLDQEGEGLIEDDKRITAGYAHBQGAIVIVNKKMD	300	
QY	301	KDSKTPKKRDEDEYKEQFLDYAOIAVSAKERFLRLFPYINDESENHKKRVOSTLN	360	
Db	301	KDSTTKKEEENRDRHQFLDYAPILFMSALTKKRRIHTLMAIIASAESHNSLRVQTNV	360	
QY	361	EYVTDASINMPPTDGRGLNVFATQVAIEPPFFVNVVNVDELHFSKRYLENOIRAA	420	
Db	361	DVIMADVANNPFTTHNGSRILKIYATQVSKPSPFVYVNDPELMHFSERLENNRIDA	420	
QY	421	FGEGTPIPIIIAKRR	435	
Db	421	FGEGTPIPIIFARAR	435	

RESULT 5
F83854
hypothetical protein BH1638 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-jun-2001
C:Accession: F83854
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: F83854
A:Status: preliminary

A: Molecule type: DNA
 A: Residues: 1-437 <STO>
 A: Cross-references: GB:AP001512; GB:BA000004; NID:J01174030; PIDN:BAB05357.1; GSPDB:G
 A: Experimental source: strain C-125
 C: Genetics:
 A: Gene: BH1538
 C: Superfamily: Mycobacterium leprae probable GRP-binding protein; translation elongat
 Query Match 74.9%; Score 1658; DB 2; Length 437;
 Best Local Similarity 70.9%; Pred. No. 1.3e-94;
 Matches 309; Conservative 67; Mismatches 60; Indels 0; Gaps 0;

[illegible]

RESULT 6
 B95199
 phosphoglycerate dehydrogenase-related protein [imported] - Streptococcus pneumoniae
 C.Species: Streptococcus pneumoniae
 C.Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C.Accession: B95199
 R.Titeltein, H.; Nelsson, K.E.; Paulsen, I.T.; Elsen, J.A.; Read, T.D.; Peterson, S.; H
 on, J.D.; Unayama, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A.Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
 A.Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A.Reference number: A95000; MUID:21357209; PMID:11463916
 A.Accession: B95199
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-436 <KUR>
 A.Cross-references: GB:A005672; PIDN:AAK75787.1; PID:g14973204; GSPDB:GN00164; TIGR:
 A.Experimental source: strain TIGR4
 C.Genetics:
 A.Gene: SPI709
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongat
 Query Match 68.8%; Score 1522; DB 2; Length 436;
 Best Local Similarity 66.2%; Pred. No. 3e-86;
 Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

Query Match 68.8%; Score 1522; DB 2; Length 436;
Best Local Similarity 66.2%; Pred. No. 3e-86;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;
QY 1 MTKPIVALGPPNKGKSTIFNRIRIGERVSYEDTPGVTQRRIYSSGEMLTTHDENIIDTG 60

```

Db 1 MALPETAIVGRPNVNGKSTLEPNRJIAGERISLIVEEGYTRDRIRYATGEMLRMSMIDTGG 60
Qy 61 IEIGDABFQOTQIRAQAEIADIDADVLIIFMVNREGLTOSDEMVAQIILYKSKREPVIAVVK 120
Db 61 IDVDVDAFPMQIKHQAEIAMEADVIVFYVSGKEGITDADEYARLKYTHKRVILAVVK 120
Qy 121 VDNKEMRTDVIYDFYSLGEGEPYPIGSGHGIGLDDLDAVVSHPGE--EEEDPYDEDTIRL 178
Db 121 VDNEMEMRNDIYDFALDGLGEPPLISSVHGICGTDVDAIENLPNEVEEENP---DVIKF 177
Qy 179 SIIGRPVNGKSSLVNALILGEDRIYVSVAQTTTDAIDTEIS-YDGGDIYULIDPAGRRKG 237
Db 178 SLIGRPVNGKSSLLNALILGEDRIYASPAQTTTDAIDTHTDGDGGEFTMIDPAGMRKG 237
Qy 238 KYVESTEKYSULRALKAIERSNVULVYDAEOGITODKKRVAGYAHOGKAVYIVYVVKMD 297
Db 238 KYVENTETKYSYMRMRARIDRSDDVYLVAINEBGIREYDKRIASFAHDAGGMITIVYVVKMD 297
Qy 298 TVEKDSKTMKRFEDVEKKEFOPLDYAQIYASAKERTRLTLEPPYINEASENHKRRVQSS 357
Db 298 TLEKDNTMTKNWEDDINEQFOYLPYAPIIFVSALITQRHLKLPDEMILKQISESONTYIPISA 357
Qy 358 TLNEVVTDAISMNPTPDKGRRLNVFYATQVALEPTEVYVNDVDELHMFHSYKRYLENOI 417
Db 358 VLNDVYMDAIAINPTPDKGRKRLIKFIYATQVAATKPTFVIFVNEEELMHPSYDRYLENOI 417
Qy 418 RAARFEGCTPIHITAKRK 435
Db 418 RKAARFEGCTPIHILAKRK 435

```


305 TMIEYEKKLEKMPFISMDPIVFISALKSQRLNKLEKVIIVQVKNNSREIKONLNDLV 364
 Db
 Oy 365 DAISMNPPTDKGRRLVNFAATOVAIEPPFFVFNVDNELMHSYKRYLENQIRAFGE 424
 | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 365 DMQTMMPL-TEFGKKLEIKRIKKTNDVPPEFLLEVNNPNIVHFSYLRYEINQIDPDET 423
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 Oy 425 GTPPIIARK 434
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 Db 424 GCPINLVKR 433

RESULT 14
 D64236
 hypothetical protein homolog MG329 - Mycoplasma genitalium
 C:Species: Mycoplasma genitalium
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
 C:Accession: D64236
 R:Frazer, C.M.; Gccayen, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
 M.: Fuhmann, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
 C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A:Title: The minimal gene complement of Mycoplasma genitalium.
 A:Reference.number: A64200; MUID:96026346; PMID:7569933
 A:Accession: D64236
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule.type: DNA
 A:Residues: 1-448 <TIGR>
 A:Cross-references: GB:U39715; GB:L43967; MID:g1046026; PID:g1046032; TIGR:MG329
 A:Experimental source: strain G-37
 C:Genetics:
 A:Genetic.code: SGC3
 A:Start.codon: GTG
 C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
 C:Keywords: duplication; GTP binding; nucleotide binding; P-loop
 F:2-121/Domain: translation elongation factor Tu homology <ET1>
 F:8-15/Region: nucleotide-binding motif A (P-loop)
 F:118-121/Region: GTP-binding NKXD motif #status atypical
 F:129-300/Domain: translation elongation factor Tu homology <ET2>
 F:185-192/Region: nucleotide-binding motif A (P-loop)
 F:297-300/Region: GTP-binding NKXD motif
 F:332-334/Region: GTP-binding SAL/K motif #status atypical

	Query Match	42.3%	Score 935.5;	DB 2;	Length 448;
	Best Local Similarity	43.2%;	Pred. No. 3.5e-50;		
	Matches 189;	Conservative 86;	Mismatches 154;	Indels 9;	Gaps 4;
Oy	6 VAIVGRNVKSTIFNRIVEGRVSIVEDPCVRDRIYSSEGMTTHDFNIITDGIEGD 65	: : :			
Db	4 VAIIGRNWVKSTIFENLIQKPMAIVSDPTPTTTDRIFGIGEMLKRIARFDITGGILAKO 63	: : :			
Oy	66 APFOQTIRAQAELIAIDEADVIIFMVNVRREGLTOSDEMYAQIIYLSKR-KPVVLAVNKVDNM 124	: : : : : : : : : : : : : : : : : :			
Db	64 TPLOQLIALGVQALISOAKAIIPLVSLIOELNSDDPFYAAKVLLKKNCKDPVILVNKAENE 123	: : : : : : : : : : : : : : : : : :			
Oy	125 EMRT---DYDIFYSLGEGEPPIPGSHGLGIAGLDLDAVVSH--FGEEEDPYDEDITRL 178	: :			
Db	124 NPKTAETELKDYSLSLGFRPVYISAHGIGTGIDMLDLVKNOLLPENND-DLAIRF 181	: :			
Oy	179 SIIRPNVKGSSLVNALIGEDRVTVSNVAGTRRAIDREYSYGDGVVLLDTACMRKKK 238	: : :			
Db	182 CVIKRPNVGSSLINOLVKORVLYSNEGSTTRAIDIPLVKNEKEKFLIDITAGIKRRGK 241	: : :			
Oy	239 VEESTERYSVLRALKAEISRNVLVIDAEGIIIEODRKRVAGVAHEOGKAVIIVNNMDT 298	: : : : : : : : : : : : : : : : : :			
Db	242 INMGIEIASYIKTKTLARSNVILLAWDGSRIPEODEVIGGLAALLPIILVNKMDL 301	: : : : : : : : : : : : : : : : : :			
Oy	299 VERASKTYKKREDEYKREFOLFDAQIAFPVASAKRTIRLRLPYLINEASENHKKRVOST 358	: : : : : : : : : : : : : : : : : :			
Db	302 VLKNNNTNAKKMKLHFHKLDEAPVLEISVLKNORLNITFEOKLIITOSLETKVATPL 361	: : : : : : : : : : : : : : : : : :			
Oy	359 LNEVVATAISMNPPTDKGRRLVNFAATOVAIEBPFFVFNVDNELMHSYKRYLENQIR 418	: : :			
Db	362 LNDVIOQOLYNOPPLRGKRRLQITVAAYORKSQIPHPFLFCNDKRVYLFHESFARLENKIR 421	: : :			

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:19:46 ; Search time 23 Seconds
(without alignments)
891.463 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MTRPIVAIVGRPNVGRKSTIF.....IRAFGEFGTPIHIARRRN 436

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt.41:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2213	100.0	436	1	ENG_A_STAM
2	1731	78.2	436	1	ENG_A_LISMO
3	1724	77.9	436	1	ENG_A_LISIN
4	1699	76.8	436	1	ENG_A_BACSU
5	1658	74.9	437	1	ENG_A_BACD
6	1528	69.0	436	1	ENG_A_STRPY
7	1523	68.8	436	1	ENG_A_STRPY
8	1522	68.8	436	1	ENG_A_STRPY
9	1520	68.7	436	1	ENG_A_STRPY
10	1251.5	56.6	439	1	ENG_A_LACIA
11	1249.5	56.5	438	1	ENG_A_THETN
12	1218.5	55.1	438	1	ENG_A_CLOPE
13	1178.5	53.3	440	1	ENG_A_CLOAB
14	1037	46.9	453	1	ENG_A_FUSNN
15	1021	46.1	452	1	ENG_A_FUSNN
16	996	45.0	449	1	ENG_A_SYNY3
17	995	45.0	442	1	ENG_A_MYCPN
18	984.5	44.5	435	1	ENG_A_MYCPN
19	935.5	42.3	448	1	ENG_A_MYCPN
20	923	41.7	490	1	ENG_A_MYCPN
21	920	41.6	490	1	ENG_A_MYCPN
22	910	41.1	490	1	ENG_A_MYCPN
23	910	41.1	490	1	ENG_A_MYCPN
24	907.5	41.0	437	1	ENG_A_MYCPN
25	906.5	41.0	437	1	ENG_A_MYCPN
26	898.5	40.6	496	1	ENG_A_MYCPN
27	896.5	40.5	495	1	ENG_A_MYCPN
28	878.5	39.7	498	1	ENG_A_MYCPN
29	877.5	39.7	465	1	ENG_A_MYCPN
30	864.5	39.1	433	1	ENG_A_MYCPN
31	862.5	39.0	519	1	ENG_A_MYCPN
32	860.5	38.9	485	1	ENG_A_MYCPN
33	858.5	38.8	485	1	ENG_A_MYCPN

34	857.5	38.7	485	1	ENG_A_MYCPN
35	856	38.7	510	1	ENG_A_MYCPN
36	852	38.5	504	1	ENG_A_MYCPN
37	846	38.2	447	1	ENG_A_MYCPN
38	842.5	38.1	439	1	ENG_A_MYCPN
39	840.5	38.0	493	1	ENG_A_MYCPN
40	808.5	36.5	479	1	ENG_A_MYCPN
41	806.5	36.4	463	1	ENG_A_MYCPN
42	797.5	36.0	483	1	ENG_A_MYCPN
43	789	35.7	476	1	ENG_A_MYCPN
44	784	35.4	438	1	ENG_A_MYCPN
45	782.5	35.4	461	1	ENG_A_MYCPN

ALIGNMENTS

RESULT 1
ID ENG_A_STAM STANDARD: PRT: 436 AA.
AC 09915:
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENG_A OR SAV1475 OR SA1307 OR MW1364.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iano J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,
RA Mutantani-Ul Y., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RT Lancet 357:1225-1240(2001).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu2;
RX MEDLINE=22040717; PubMed=1204378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-acquired MRSA."
RT Lancet 359:1819-1827(2002).
RL -1- SIMILARITY: BELONGS TO THE ERA/TRE FAMILY OF GTP-BINDING PROTEINS. ENG_A SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: AP003362; BAB57637.1; -
CC EMBL: AP003134; BAB42568.1; -
CC EMBL: AP004827; BAB95229.1; -
CC PIR: C89926; C89926.
CC HAMAP: MF_00195; -; 1.
CC InterPro: IPR005289; GTP-binding_dom.
CC InterPro: IPR006073; GTP1_OBG.

```

DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRfams: TIGR00650; MG442; 2.
DR TIGRfams: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 GTP 1 (POTENTIAL).
FT NP_BIND 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 GTP 1 (POTENTIAL).
FT NP_BIND 119 GTP 1 (POTENTIAL).
FT NP_BIND 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 GTP 2 (POTENTIAL).
FT NP_BIND 229 GTP 2 (POTENTIAL).
FT NP_BIND 293 GTP 2 (POTENTIAL).
FT NP_BIND 294 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 48979 MW; 9EA4281B62F229D3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 2213; DB 1; Length 436;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKPIVAIVGRPNVSGSTIFNRIVGERVSIVEDPGVTRDRIRYSGEWLTHDFNIIDTGG 60
DB 1 MTKPIVAIVGRPNVSGSTIFNRIVGERVSIVEDPGVTRDRIRYSGEWLTHDFNIIDTGG 60
QY 61 IETGAPFQTOIRAQAEIAIDADVITFVWVNRREGLTOSDEWVAQILYKSKRPVLAIVK 120
DB 61 IETGAPFQTOIRAQAEIAIDADVITFVWVNRREGLTOSDEWVAQILYKSKRPVLAIVK 120
QY 121 VDNMEMRTDVFYSISGFEPPYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
DB 121 VDNMEMRTDVFYSISGFEPPYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
QY 121 VDNMEMRTDVFYSISGFEPPYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
DB 121 VDNMEMRTDVFYSISGFEPPYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
QY 181 ITRPNVGSLLVNAIIGEDRVIVSNVAGTTRADITFYSIDODVYLITAGMKKKGYK 240
DB 181 ITRPNVGSLLVNAIIGEDRVIVSNVAGTTRADITFYSIDODVYLITAGMKKKGYK 240
QY 241 ESTEKSVLRALKATERSNVVLVIDAEGGITEODKRVAGVHOGKAVIYVNMKDTVE 300
DB 241 ESTEKSVLRALKATERSNVVLVIDAEGGITEODKRVAGVHOGKAVIYVNMKDTVE 300
QY 301 KDSKTMKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPIYNEASENHKKRVOSTLN 360
DB 301 KDSKTMKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPIYNEASENHKKRVOSTLN 360
QY 361 EYVTDASINNPPTDKGRNLNVFYATQVAIEPPTVFVNDVLMHFSKRYLENOIRAA 420
DB 361 EYVTDASINNPPTDKGRNLNVFYATQVAIEPPTVFVNDVLMHFSKRYLENOIRAA 420
QY 421 FGEEGPTIHIIARKRN 436
DB 421 FGEEGPTIHIIARKRN 436

RESULT 2
ENGA_LISMO STANDARD; PRT; 436 AA.
AC 08158;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein enga.
GN ENGA OR LMO1937.
OC Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SGD-se / Setovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baguerio F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charif A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA Entlan K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

```

```

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Klett J., Kuhn M., Kunst F., Kurapkai G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomes of Listeria species."
RL Science 294:849-852(2001).
CC -i- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL591981; CAD00015.1; -.
CC PIR: A11316; A11316.
CC Listlist: LMO01937; -.
CC HAMAP: ME_00195; -.
CC DR InterPro: IPR005289; GTP-binding_dom.
CC DR InterPro: IPR006073; GTP_OBG.
CC DR InterPro: IPR002917; MMR_HSR1.
CC DR Pfam: PF01926; MMR_HSR1; 1.
CC DR PRINTS: PR00326; GTP_OBG.
CC DR TIGRfams: TIGR00650; MG442; 2.
CC TIGRfams: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 GTP 1 (POTENTIAL).
FT NP_BIND 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 GTP 1 (POTENTIAL).
FT NP_BIND 119 GTP 1 (POTENTIAL).
FT NP_BIND 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 GTP 2 (POTENTIAL).
FT NP_BIND 229 GTP 2 (POTENTIAL).
FT NP_BIND 293 GTP 2 (POTENTIAL).
FT NP_BIND 294 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 49144 MW; 5B435C8701AE6A3 CRC64;

Query Match
Best Local Similarity 78.2%; Score 1731; DB 1; Length 436;
Matches 320; Conservative 65; Mismatches 50; Indels 0; Gaps 0;

QY 1 MTKPIVAIVGRPNVSGSTIFNRIVGERVSIVEDPGVTRDRIRYSGEWLTHDFNIIDTGG 60
DB 1 MTKPIVAIVGRPNVSGSTIFNRIVGERVSIVEDPGVTRDRIRYSGEWLTHDFNIIDTGG 60
QY 61 IETGAPFQTOIRAQAEIAIDADVITFVWVNRREGLTOSDEWVAQILYKSKRPVLAIVK 120
DB 61 IDSDEPLEQIRAQAEIAIDADVITFVWVNRREGLTOSDEWVAQILYKSKRPVLAIVK 120
QY 121 VDNMEMRTDVFYSISGFEPPYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
DB 121 VDNMEMRTDVFYSISGFEPPYPISSGSHGLGDLIDAVVSHFGEEDPDYEDTIRLSI 180
QY 181 ITRPNVGSLLVNAIIGEDRVIVSNVAGTTRADITFYSIDODVYLITAGMKKKGYK 240
DB 181 ITRPNVGSLLVNAIIGEDRVIVSNVAGTTRADITFYSIDODVYLITAGMKKKGYK 240
QY 241 ESTEKSVLRALKATERSNVVLVIDAEGGITEODKRVAGVHOGKAVIYVNMKDTVE 300
DB 241 ESTEKSVLRALKATERSNVVLVIDAEGGITEODKRVAGVHOGKAVIYVNMKDTVE 300
QY 301 KDSKTMKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPIYNEASENHKKRVOSTLN 360
DB 301 KDSKTMKFEDEVRKEFOFLDYAQIAFVSAKERTRLTLFPIYNEASENHKKRVOSTLN 360
QY 361 EYVTDASINNPPTDKGRNLNVFYATQVAIEPPTVFVNDVLMHFSKRYLENOIRAA 420
DB 361 DIYSDAVANPSPMDKGRILKIFYTQVAVKPTFVFNDELHMFYSERLENRIRGA 420
QY 421 FGEEGPTIHIIARKRN 435

```

```

Db      421  FPFEGTPPIVIAKRK 435

RESULT 3
ENGA_LISIN
ID      ENGA_LISIN      STANDARD:      PRT:      436 AA.
AC      Q92A71;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Probable GTP-binding protein enga.
GN      ENGA OR LIN2051.
OS      Listeria innocua.
OC      Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX      NCBI_TaxID=1642;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-CLIP 11262 / Serovar 66;
RX      MEDLINE=21537279; PubMed=11679669;
RA      Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA      Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA      Chabdit A., Chetouani F., Cove E., de Daruvar A., Dehoux P.,
RA      Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA      Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA      Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA      Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkhat G.,
RA      Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA      Nordstak G., Novella S., de Pablos B., Perez-Diaz J.-C., Puceat R.,
RA      Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA      Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RT      "Comparative genomics of Listeria species.";
RL      Science 294:849-852(2001).
CC      -1- SIMILARITY: BELONGS TO THE ERA/TRNE FAMILY OF GTP-BINDING
CC      PROTEINS. ENGA SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: AL596170; GAC97281.1; -
DR      PIR: A11688; A11688
DR      ListList: LIN2051; -
DR      HAMAP: MF_00195; -; 1.
DR      InterPro: IPR005289; GTP-binding_dom.
DR      InterPro: IPR006073; GTP_OBG.
DR      InterPro: IPR002917; MMR_HSR1.
DR      InterPro: IPR005225; Small_GTP.
DR      Pfam: PF01926; MMR_HSR1. 1.
DR      PRINTS: PR00326; GTP_OBG.
DR      TIGRFAMs: TIGR00650; MG442. 2.
DR      TIGRFAMs: TIGR00231; small_GTP. 2.
KW      GTP-binding; Repeat; Complete proteome.
FT      NP_BIND      10      17      GTP 1 (POTENTIAL).
FT      NP_BIND      57      61      GTP 1 (POTENTIAL).
FT      NP_BIND      119     122     GTP 1 (POTENTIAL).
FT      NP_BIND      182     189     GTP 2 (POTENTIAL).
FT      NP_BIND      229     233     GTP 2 (POTENTIAL).
FT      NP_BIND      294     297     GTP 2 (POTENTIAL).
FT      NP_BIND      436 AA; 49144 MW; 1A435C1970EB66C8 CRC64;
SQ      SEQUENCE

```

```

Query Match      77.9%; Score 1724; DB 1; Length 436;
Best Local Similarity 73.1%; Pred. No. 2e-98;
Matches 318; Conservative 67; Mismatch 50; Indels 0; Gaps 0;

```

```

OY      1 MKRPVATVGRNPKSTFENRIVGEVSIVEDTQCVTRDRYSSEWMLTHFNITDGG 60
DB      1 MAKPVAVTGRPNVGKSTFENRIVGEVSIVEDVPGVTRDRITNSAEWLKGFENIIDTGG 60

```

```

OY      61 IEIGDAFQOTQIRAQAEIAIDEDAVIIFMVNVNREGLTOSDEKVAQILKSKKPVAVNK 120
DB      61 IDLSDEPFLEQIRAAQAEIAIDEDAVIIFITNREGVTDADQVAKILVRSNKPVLAINK 120
OY      121 VDNMEKRDVDYDFYSLGCEPPIPGISGICLGIDLDDNVSHFSGEEDRPYEDTIRLSI 180
DB      121 VDNPEMRQOYIDFYSLGCEPPIPGISGICLGIDLDDNVRAHFPKEEEDRPDETVKPSL 180
OY      181 IGRPNVGSLSVALNALLGEDRVVSNVAGTTRPAIDTEXSYDGODVLLDTAGMRKKGKY 240
DB      181 IGRPNVGSLSVALNALLGEDRVVSDIAGTTRAITDTTTFDQODVYMDIAGMRKRGKY 240
OY      241 ESTERYSVLRALKAIERSNVVLVIDAEGCIEQDKRVAAGAHQEGKAVYIVNKNWTV 300
DB      241 ESTERYSVLRAMRAIERSDVLVIVNAEGRIGQDKRIAGYAHDAGRALITIVNKNWDAIN 300
OY      301 KDSKTMKRFEDVREKPEQFLDYAQAIPASAKRTLRITLPIFINASEMHNKRVQSTLN 360
DB      301 KDEKTNWTEIDIRQFQFLSYAPIVFVSAKTRKRLNLEPLINQVSDNHSILRVQSSMLN 360
OY      361 EYVTAISNPTPTQKGRNLNVEYATQVAVIEPPTFVFNVDVLMHFYKRYLEQIRAA 420
DB      361 DVISDAVAMNPSPMKGRKRLKIFTTQVAVKPTFVFNVDVLMHFYERPLENRIRRA 420
OY      421 FPFEGTPPIIARIKR 435
DB      421 FPFEGTPPIVIAKRK 435

```

RESULT 4

```

ENGA_BACSU
ID      ENGA_BACSU      STANDARD:      PRT:      436 AA.
AC      P50743;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Probable GTP-binding protein enga.
GN      ENGA.
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168 / Marburg;
RX      MEDLINE=96349105; PubMed=8760912;
RA      Sorokin A.V., Azavedo V., Zumbstein E., Galleron N., Enhlisch S.D.,
RA      Serror P.;
RT      "Sequence analysis of the Bacillus subtilis chromosome region between
RT      the serA and kds loci cloned in a yeast artificial chromosome.";
RL      Microbiology 142:2005-2016(1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=98044033; PubMed=9384377;
RA      Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA      Azavedo V., Bertsche M.G., Bessieres P., Bolotin A., Borchert S.,
RA      Boriss R., Bourasier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA      Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,
RA      Choi S.K., Codani J.J., Connerton I.F., Cumings N.J., Daniel R.A.,
RA      Denzot F., Devine K.M., Dusterhoft A., Enhlisch S.D., Emmerman P.T.,
RA      Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA      Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA      Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA      Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut L.,
RA      Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA      Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA      Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
RA      Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA      Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA      Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA      Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA      Parro V., Pohl T.M., Portetelie D., Porwollik S., Prescott A.M.,
RA      Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

```

```

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield E.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terstra P., Togononi K.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarotti A.,
RA Viari A., Wambolt R., Wedler E., Wedler H., Wetzenecker T.,
RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE OF 185-436 FROM N.A.
RC STRAIN-168;
RA MEDLINE=96011379; PubMed=7592341;
RA Moridoni H.R., de Mendoza D., Cronan J.E. Jr.;
RT "Synthesis of sn-glycerol 3-phosphate, a key precursor of membrane
RT lipids, in Bacillus subtilis."
RL J. Bacteriol. 177:5899-5905(1995).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: L47648; AAC83966.1; -
DR EMBL: U32164; CAB14200.1; -
DR PIR: A69936; A69936.
DR Subtilisin; BG11443; engA.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small-GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRfams: TIGR00650; MG442; 2.
DR TIGRfams: TIGR00231; small_GTP; 2.
KM GTP-binding: Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 189 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
FT CONFLICT 186 187 VG -> CR (in Ref. 2).
SQ SEQUENCE 436 AA; 48769 MM; A5CC7028F8B5A442 CRC64;

```

```

Query Match 76.8%; Score 1699; DB 1; Length 436;
Best Local Similarity 73.6%; Pred. No. 6.8e-97;
Matches 320; Conservative 57; Mismatches 58; Indels 0; Gaps 0;

```

```

QY 181 IGRPNWKKSLVNAIIGEDRYVSNAGTTTROIIDREYSYDGDVYLITAGMKKKGY 240
DB 181 IGRPNWKKSLVNAIIGEDRYVSNAGTTTROIIDREYSYDGDVYLITAGMKKKGY 240
QY 241 ESTREKSYVALKAIRSNVNLVVADEGIIKODRVRVGYAEHOGKAVIYNKMDTVE 300
DB 241 ESTREKSYVALKAIRSNVNLVVADEGIIKODRVRVGYAEHOGKAVIYNKMDTVE 300
QY 301 KDSRTKKKEDEVRKKEFQFLDYAQAIFVSAKERTRLTLEPTINEASENKKRQSSSTLN 360
DB 301 KDSRTKKKEDEVRKKEFQFLDYAQAIFVSAKERTRLTLEPTINEASENKKRQSSSTLN 360
QY 361 EYVTDALISMPPTTDGRLNVTYATQVAIEPTFPVFNVDVLMFSTKRYLENIIRA 420
DB 361 DVTMDAVANMPPTTHNGSRKTYATQVSVPSPFVFNVDVLMFSTKRYLENIIRA 420
QY 421 FGEGTPIHIAKRK 435
DB 421 FGEGTPIHIAKRK 435

RESULT 5
ENG_A_BACHD STANDARD; PRT; 437 AA.
ID ENG_A_BACHD
AC Q9KCD4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENG_A_OR_BH1638
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=86665;
RX NCBI_TaxID=86665;
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AP001512; BAB05357.1; -
DR PIR: F83854; F83854.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small-GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRfams: TIGR00650; MG442; 2.
DR TIGRfams: TIGR00231; small_GTP; 2.
KM GTP-binding: Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 182 189 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 437 AA; 49024 MM; C1A48D55A39AD4B CRC64;

```

Query	March	74.9%	Score 1658	DB 1	Length 437
Best Local Similarity	70.9%	Pred. No. 2.2e-94			
Matches	309	Conservative	67	Mismatches	60
				Indels	0
				Gaps	0
QY	1	MTKPIVAIVGRPNVNGSKSTIFNRIVIGERSIYEDPPGVTRDRIRYSSGSEWLTDFDNIIDGG	60		
Db	1	MSKPIVAIVGRPNVNGSKSTIFNRIVIGERSIYEDPPGVTRDRIRYSSGSEWLTDFDNIIDGG	60		
QY	61	IEIGDAFPOQIIRQAQELAIIDEAIVIFWVWVREGSLTQSDENWAOILTKSKPPVLAAYK	120		
Db	61	IEIGDEPLLVQMRQAQELAIKEADVIIFIVNGREGVTAADQEVAKILFRSKPPVLAAYK	120		
QY	121	VDNMERTDVIYDFSLGCEPEYPISSGSLGLDGLDLVWVSHFGEEDDPYDEDTIRLSI	180		
Db	121	IDHDMQBELTEFSLGIGDPIPSGANGSLGLDLDCVHFPEDEGDDDEDTIRLSI	180		
QY	181	IGRPVNGKSSLVNAILGEDRYIVSNVAGTTBDADIDTEXSYDGQDVLIDTAGMRKKGVY	240		
Db	181	IGRPVNGKSSLVNAILGGERIVYISNPCTTRDAIDTAFSRDDQEVLLIDTAGMRKKGVY	240		
QY	241	ESTPKYSVLRLAKRIERSNVLYVYIDAQGIIEODKRAAGTAHEDGKAVVIVYKWDIYE	300		
Db	241	ESTPKYSVLRLAKRIERSDVVLYVNGEGEIIEDDKTAGAHEGRAIIIVYKWDIAVE	300		
QY	301	KDSITKMKFEDEYVKEFEFLDYADIAFASAKERPLTLPFYINASHNKKRVQSSLTN	360		
Db	301	KDKTTLRFOOKIDFEDFEDYADVLYFVSATKTORLQVLYLAVKVSENNHLRPTHTVLN	360		
QY	361	EVYTDIAISMNPTPDKGRNLNVFYATOVAIEPPTFVVEVNDVELMHFSYKRYLENOIRAA	420		
Db	361	LVYDVAAMNPTPDHGRKRLKINVYTVQAVGPRPFVFNDELMHFSYARLENRIDT	420		
QY	421	FGEGTPIHITARRKN 436			
Db	421	FEFGTPIKITARRKN 436			

RESULT 6

ENGA_STRPY

ID ENGA_STRPY STANDARD: PRI: 436 AA.

Q9A1D0

28-FEB-2003 (Rel. 41, Created)

28-FEB-2003 (Rel. 41, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Probable GTP-binding protein enga

ENGA OR PGDA OR SPY0341 OR SPM18_0334.

OS Streptococcus pyogenes, and

OS Streptococcus pyogenes (serotype M18).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

NCBI_TaxID=1314, 186103;

[1]

SEQUENCE FROM N.A.

STRAIN=SE370 / ATCC 700294 / Serotype M1;

MEDLINE=21192684; PubMed=11296296;

RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Seate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian X., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;

RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";

RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

[2]

SEQUENCE FROM N.A.

STRAIN=MGAS8232 / Serotype M18;

MEDLINE=21927593; PubMed=11917108;

RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S., Sylla G.L., Sturdevant D.E., Ricklets S.M., Porcella S.F., Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q., Kapur V., Daly J.A., Veasy L.G., Musser J.M.;

RA "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.";

RT

RL	Proc. Natl. Acad. Sci. U.S.A.	99:4668-4673(2002).
CC	-1-	SIMILARITY: BELONGS TO THE ENA/TRME FAMILY OF GTP-BINDING PROTEINS. ENGA SUBFAMILY.
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----	
CC	EMBL: AEO06498; AAC33393.1; -.	
DR	EMBL: AEO09978; AAL97088.1; -.	
DR	HAMAP: ME_00195; -. 1.	
DR	InterPro: IPR005289; GTP-binding_dom.	
DR	InterPro: IPR006073; GTP1_OBG.	
DR	InterPro: IPR002917; MMR_HSR1.	
DR	InterPro: IPR005225; Small_GTP.	
DR	Pfam: PF01926; MMR_HSR1. 1.	
DR	PRINTS: PR00326; GTPLOBG.	
DR	TIGRFAMS: TIGR00650; MG442; 2.	
KM	TIGRFAMS: TIGR00231; small_gtp; 2.	
FT	GTP-binding; Repeat. Complete proteome.	
FT	NP_BIND 10 17 GTP 1 (POTENTIAL).	
FT	NP_BIND 57 61 GTP 1 (POTENTIAL).	
FT	NP_BIND 119 122 GTP 1 (POTENTIAL).	
FT	NP_BIND 181 188 GTP 2 (POTENTIAL).	
FT	NP_BIND 229 233 GTP 2 (POTENTIAL).	
FT	NP_BIND 294 297 GTP 2 (POTENTIAL).	
SO	SEQUENCE 436 AA; 48801 MW; ESDDED7D51675322AA CRC64;	
Query Match	Best Local Similarity 69.0%; Score 1528; DB 1; Length 436;	
Matches 288;	Conservative 75; Pred. No. 1.9e-86;	
	Mismatches 71; Indels 2; Gaps 2;	
QY	1 MTKPIVAIVGRPNVGKSTINRNIRIGEVSIVEDTPGVTTRDRITSSGEMLTIDFNII DTGG 60	
DB	1 MVLPETVAIVGRPNVGKSTLFENRISIERVEDEGVTRDIRIATGEMLNROFSLLDTGG 60	
QY	61 IEIGDAPOFOQIRAOAEIAIDEADVIIFMVNVREGLOSDEMAVOILYKSKKPVLAVNK 120	
DB	61 IDVDYDAFFMEDIKIQAOIAMEDADVIFYVSGKGCVTDADAEYSKILKRTITPVILAVNK 120	
QY	121 VDNNEMRTDYVDFYSLGFGEPYPYPSISGHGLGDLDAVVSHFGEEDDPYDEDTIRLSI 180	
DB	121 VDNPEMNNDIYDFSLGLGDPPYPSVSHYGIGTGVDLAIVENLVPEAEEND-DIIRSL 179	
QY	181 IGRPNVGKSSLVNAILGEDRYVSNVAGTTRDADITEXS-YDGDDYVLIDTAGKRKKGV 239	
DB	180 IGRPNVGKSSLINAILGEDRVIASPVAGTTRDADITHTFDADGCEFTMDIPAGMKRSKI 239	
QY	240 YESREKSVLRALKAIERSNVVLAVIDAEOGIIEODRKRVAGVAHEOGKAVYIVNKMDTV 299	
DB	240 YEINTEKSYVMKAMKAIIRDSDVVLVINABEIRIEDRIKRIAGFAHBAKGMTIIVNKMDTI 299	
QY	300 EKDSKTAKKKEFEDEVKREFOFDLYAQIAFVSAKERTRLTLLEPYINEASENHKKRVOSSL 359	
DB	300 DKDHHTYAKKNADRIDRQFOFLTYAPIIFVSLATRGQRUNKLPDLCLKRISESQNKRIIPSAVL 359	
QY	360 NEVVYTDASINMPPTDKGRNLVEYATQAVALIEPTFPVVFVNVDVELMHFSYKRYLENQIRA 419	
DB	360 NDVIMDAIAINPPTDKGRKLIFYAQOVSAVKPTFPVVFVNEELMHFSYLDRLENOIRA 419	
QY	420 AFGEQGPIHIARKR 435	
DB	420 AFTEFGPIHLIARRK 435	
RESULT 7		
ENGAL_STRP3	STANDARD:	PRT: 436 AA.
ID ENGAL_STRP3		
CD QNRKJ8;		

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENGA OR PGD OR SPYM3.0249 OR SP51610.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=1984466;
RX STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=2113808; PubMed=1212206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayaashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes: a comparative analysis
RT of S. pyogenes SSI-1, SPI70 and MGAS8232."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AEO14141; AAM78856.1; -
DR EMBL: AP005146; BAC64705.1; -
DR HAMAP: MF_00195; -; 1.
DR Interpro: IPR005289; GTP-binding_dom.
DR Interpro: IPR006073; GTP1_OBG.
DR Interpro: IPR002917; MMR_HSR1.
DR Interpro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP1_OBG.
DR TIGRFAMs: TIGR00650; MG442; 2.
DR TIGRFAMs: TIGR00231; small_GTP; 2.
KM GTP-binding: Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 181 188 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 48771 MW; F2DFEC5B674323A CRC64;
Query Match 68.88; Score 1523; DB 1; Length 436;
Best Local Similarly 65.88; Pred. No. 3.9e-86;
Matches 287; Conservative 75; Mismatches 72; Indels 2; Gaps 2;

```

```

QY 121 VDNEMKRTDYDYSIGFCPPYISGSHGICGLDLDAVYSHFCEEDPDYEDTIRLSI 180
DB 121 VDNEMKNDIYDYSIGLGPVSSVHGIGTDVDAIYENLPVEAEEND-DIIRFSL 179
QY 181 IGRPNVGSSTLVNATIGEDVIVSNVAGTTRDAIDREYS-YDQODVLDTAGMRKGV 239
DB 180 IGRPNVGSSTLVNATIGEDVIVSNVAGTTRDAIDREYS-YDQODVLDTAGMRKGV 239
QY 240 YESTEKYSVALKATERSNVVLVIDAEGIIEDOKRVAGVAHEQGAIVVKNMDTV 299
DB 240 YENTEKYSVARARNAIRDROVYLMVINEEGINEYDKRIGAFHAGKMIIVNKMDAI 299
QY 300 EKDSKTKKPEDEYKKEFQFLDYAQAFAFSAKERTRLTLPYINAESENHKKRVOSTL 359
DB 300 DKDNHVTAKMEADIRDOFQFLTYAPITFVSALTORKLPLDLIKRISESONKRIPSAVL 359
QY 360 NEVYTAISNPNPTQGRRLNPFYATQVAIEPTFPVFNVDVLMHFSKRYLENDIRA 419
DB 360 NDVIMDAIAPPTDGRKLKIFATQVSKPPTFVNEBELMHFSLRLENDIRA 419
QY 420 AFGEFGPIHIIARRK 435
DB 420 AFTEFGPIHIIARRK 435
RESULT 8
ENGA_STRPN
ID ENGA_STRPN STANDARD; PRT; 436 AA.
AC 097PC9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable GTP-binding protein engA.
GN ENGA OR SERA OR SPI709 OR SPI1553.
OS Streptococcus pneumoniae, and
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Ufferbach T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborno W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W.G., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McHenry S.M., McHenry M., Mclester K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.T.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6."
RL J. Bacteriol. 183:5709-5717(2001).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR1 EMBL: AE007464; AAK75787.1; -
DR EMBL: AE008523; AAL00357.1; -
DR PIR: B95199; B95199.
DR PIR: H98065; H98065.
DR TIGR: SP1709; -
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRFAMS: TIGR00650; MG442; 2.
DR TIGRFAMS: TIGR00231; small_GTP; 2.
KM GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 181 188 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 49082 MW; B013890D5285BBD CRC64;

Query Match 68.8%; Score 1522; DB 1; Length 436;
Best Local Similarity 66.2%; Pred. No. 4.4e-86;
Matches 290; Conservative 73; Mismatches 69; Indels 6; Gaps 3;

QY 1 MTKPIVAIVGRPNNGKSTIFNRIYGERYSIYEDPGVTRDRYSGEWLTHDENTIDGG 60
DB 1 MALPTIAIVGRPNNGKSTIFNRIYGERYSIYEDPGVTRDRYSGEWLTHDENTIDGG 60
QY 61 IEIGDAPFOTOIRAOAETIAIDADYIIFPMVNRGLTQSDPMVAOILKSKRPVLAIVNK 120
DB 61 IDVDAPFMEIKIKOAEIAMEADYIVVSGKEGIDDADEVARKLTKTKRPVLAIVNK 120
QY 121 VDNEMKRTDYDFYSLGFGEPYPIGSGHGLGLDLDVAVSHFGE--EEDDPYEDTIRL 178
DB 121 VDNEMKRTDYDFYSLGFGEPYPIGSGHGLGLDLDVAVSHFGE--EEDDPYEDTIRL 178
QY 179 SIIGRPVNGKSLVNAIIGEDRYVSNVAGTTRDAIDTEYS-YDGQDYVLIDTAGMRKG 237
DB 178 SLIGRPVNGKSLVNAIIGEDRYVSNVAGTTRDAIDTEYS-YDGQDYVLIDTAGMRKG 237
QY 238 KVESTEKYSVLRAIKAIERSNVLVVDAEOGIIEDOKRVAGYAHEGKAVVIVNKKMD 297
DB 238 KVESTEKYSVLRAIKAIERSNVLVVDAEOGIIEDOKRVAGYAHEGKAVVIVNKKMD 297
QY 298 TLEKDNHMKWEEDIREQFQYLPAPRIIFVSALTKQHLKLEIKIISSONRIPISA 357
DB 298 TLEKDNHMKWEEDIREQFQYLPAPRIIFVSALTKQHLKLEIKIISSONRIPISA 357
QY 358 TLNEVYDAISMNPTPDGKRLNVFATQVAIEBPFTVNVVDVLEMHESYKRYLENOI 417
DB 358 VLNVDVDAISMNPTPDGKRLNVFATQVAIEBPFTVNVVDVLEMHESYKRYLENOI 417
QY 418 RAAGFEGEPTIHIAARR 435
DB 418 RAAGFEGEPTIHIAARR 435

RESULT 9
ID ENGA_LACIA STANDARD: PRT: 436 AA.
AC Q9CHH6: 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein enga.
GN ENGA OR YPHL OR IL0755.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;

RA Boletín A., Wincker P., Manger S., Jallion O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RA "The complete genome sequence of the lactic acid bacterium Lactococcus
RA lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -I- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR1 EMBL: AE006309; AAK04853.1; -
DR PIR: C86719; C86719.
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP_OBG.
DR TIGRFAMS: TIGR00650; MG442; 2.
DR TIGRFAMS: TIGR00231; small_GTP; 2.
KM GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 119 122 GTP 1 (POTENTIAL).
FT NP_BIND 181 188 GTP 2 (POTENTIAL).
FT NP_BIND 229 233 GTP 2 (POTENTIAL).
FT NP_BIND 294 297 GTP 2 (POTENTIAL).
SQ SEQUENCE 436 AA; 48866 MW; 1D9D0081E9865B9D CRC64;

Query Match 68.7%; Score 1520; DB 1; Length 436;
Best Local Similarity 66.3%; Pred. No. 5.9e-86;
Matches 289; Conservative 69; Mismatches 76; Indels 2; Gaps 2;

QY 1 MTKPIVAIVGRPNNGKSTIFNRIYGERYSIYEDPGVTRDRYSGEWLTHDENTIDGG 60
DB 1 MSLPTVAIVGRPNNGKSTIFNRIYGERYSIYEDPGVTRDRYSGEWLTHDENTIDGG 60
QY 61 IEIGDAPFOTOIRAOAETIAIDADYIIFPMVNRGLTQSDPMVAOILKSKRPVLAIVNK 120
DB 61 IELSDPEPTMELIRAOAETIAIDADYIIFPMVNRGLTQSDPMVAOILKSKRPVLAIVNK 120
QY 121 VDNEMKRTDYDFYSLGFGEPYPIGSGHGLGLDLDVAVSHFGEEDDPYEDTIRLSI 180
DB 121 VDNEMKRTDYDFYSLGFGEPYPIGSGHGLGLDLDVAVSHFGEEDDPYEDTIRLSI 180
QY 181 IGRPNVGSLSLNAIIGEDRYVSNVAGTTRDAIDTEYS-YDGQDYVLIDTAGMRKGK 239
DB 180 IGRPNVGSLSLNAIIGEDRYVSNVAGTTRDAIDTEYS-YDGQDYVLIDTAGMRKGK 239
QY 240 YESTEKYSVLRAIKAIERSNVLVVDAEOGIIEDOKRVAGYAHEGKAVVIVNKKMDTV 299
DB 240 YENFEKYSVNRAMRAIDRSIDVLYVNAEGRREYDKRIAGFAHAGKGLIVVKKMPTL 299
QY 300 EKDSKTKMKFEDEVAKKEFQPLDYAOIAFVSAKERTRLTLPYINASENHKKRVOSTL 359
DB 300 EKDNMTKMFLEIETKKEFLDYAPIVYVSAKTGRLKMDIMEIHHAONLRISSVL 359

```

QY 360 NEVYTDALSMNPPPTDKGRRLNVFAVNOVAIEPPFVFNVDVLMHFSYKRYLENOIRA 419
DB 360 NVIMDAVAINPTPTDKGRRLKIFYATQAIKRPPTFVFNVEBELMHFSYRLLENQIRK 419
QY 420 AFGFESTPIHIIARKR 435
DB 420 AFEFESTPVHLIARKR 435

RESULT 10
ENGA_THETN
ID ENGA_THETN STANDARD: PRT: 439 AA.
AC OGRSUL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENGA OR TTEL619.
OS Thermoaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;
OC Thermoaerobacteriaceae; Thermoaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE013117; AAM24821.1; -
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR SMART: SM00382; AAA; 2.
DR TIGRFRAMS: TIGR00650; MG442; 2.
DR TIGRFRAMS: TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 GTP 1 (POTENTIAL).
FT NP_BIND 57 GTP 1 (POTENTIAL).
FT NP_BIND 120 GTP 1 (POTENTIAL).
FT NP_BIND 183 GTP 2 (POTENTIAL).
FT NP_BIND 230 GTP 2 (POTENTIAL).
FT NP_BIND 295 GTP 2 (POTENTIAL).
FT NP_BIND 298 GTP 2 (POTENTIAL).
SQ SEQUENCE 439 AA: 49432 MW: 7705252R0F268DCE5 CRC64;

Query Match 56.6%; Score 1251.5; DB 1: Length 439,
Best Local Similarity 54.1%; Pred. No. 1,5e-69;
Matches 237; Conservative 92; Mismatches 104; Indels 5; Gaps 4;

```

```

QY 119 NKVDNM-EKRTDYDFYSLGFEPPPISSGHGIGLDLDAVSHGEEEDPDYEDTIR 177
DB 120 NKVDSKEMDPASPYDFDKIGLGEPIPISSANGIGIGELLDVEYIKRLPENDVE-YEEETIK 178
QY 178 LSTIGRPNNKSSLVNAILGEDRVYISNAGTTRDAIDTETYSIDGQDYVLDTAGKRRKG 237
DB 179 IAVIGRPNNKSSLVNRIIGEEERYIVSDIPGTRDAIDPFPKDGKNYLLIDAGIRRSK 238
QY 238 KYVESTKYSVLKALAIERSNVLVIAEOGIIODKRVAGYAEQCKAVYIVYVKNKD 297
DB 239 RISESEIERISVLRALAIERADICLMDTBEPTQDTKIACIYAEENGKGIITLVNKKD 298
QY 298 IYKDSKTKMKFEDEYRKKEPFLDYAQIAFVSAKERTRLTLEPYINEASENHKKRVQSS 357
DB 299 IYKDSNTFYKEYTEMIREKIAFISFAPILFISAKTGQRHKKVLETVDKWEEYKRRITTG 358
QY 358 TLNEVYTDALSMNPPPTDKGRRLNVFAVNOVAIEPPFVFNVDVLMHFSYKRYLENOI 417
DB 359 LNNVNLNEMAMLEPPSSKGRPIKIYATQVGTTRPFTVIFVNEPELLHFSYRPLENTI 418
QY 418 RAAFGEFGPIHIIARKR 435
DB 419 RQNFGEFVPIVISTKKR 436

RESULT 11
ENGA_CLOPE
ID ENGA_CLOPE STANDARD: PRT: 438 AA.
AC OGRXUL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engA.
GN ENGA OR CPEI755.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohnishi K., Hirakawa H.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
CC PROTEINS. ENGA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP003191; BAB81461.1; -
DR HAMAP: MF_00195; -; 1.
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR006073; GTP1_OBG.
DR InterPro: IPR002917; MMR_HSR1.
DR InterPro: IPR001806; Ras_trsfmng.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF01926; MMR_HSR1; 1.
DR PRINTS: PR00326; GTP1OGB.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00382; AAA; 2.
DR TIGRFRAMS: TIGR00650; MG442; 2.
DR TIGRFRAMS: TIGR00231; small_GTP; 2.

```

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements/ or send an email to license@ebi.ac.uk).
CC	-----
DR	EMBL; AE007680; AAK79677.1; -.
DR	PIR; B97111; B97111.
DR	HAMAP; MF_00195; -. 1.
DR	InterPro; IPR005289; GTP-binding_dom.
DR	InterPro; IPR006073; GTP1_OBG.
DR	InterPro; IPR002917; MMR_HSR1.
DR	InterPro; IPR001806; Ras_transfmrng.
DR	InterPro; IPR005225; Small_GTP.
DR	Pfam; PF01926; MMR_HSR1; 1.
DR	PRINTS; PR00326; CTP1OBG.
DR	PRINTS; PR00449; RASTRNSFRANG.
DR	TIGRFAMS; TIGR00650; MG442; 2.
DR	TIGRFAMS; TIGR00231; small_gtp; 2.
KW	GTP-binding; Repeat; complete proteome.
FT	NP_BIND 10 17 GTP 1 (POTENTIAL).
FT	NP_BIND 57 61 GTP 1 (POTENTIAL).
FT	NP_BIND 120 123 GTP 1 (POTENTIAL).
FT	NP_BIND 183 190 GTP 2 (POTENTIAL).
FT	NP_BIND 230 234 GTP 2 (POTENTIAL).
FT	NP_BIND 295 298 GTP 2 (POTENTIAL).
SQ	SEQUENCE 438 AA; 49797 MW; 7765755BD7355C67 CRC64;
Query Match	55.1%; Score 1218.5; DB 1; Length 438;
Best Local Similarity	52.8%; Pred. No. 1.5e-67;
Matches 230; Conservative	88; Mismatches 117; Indels 1; Gaps 1;
QY	1 MTRPIVALVGRPNVGKSTIFNRIVGERVSYVEDPGVTQRORIYSGGMLTHDFNIIDTG 60
DB	1 MARPLVIIVRPNPKGSTLFNLAKGRVSIEDPGVTRORIAVESEWVGKKPTIIDTG 60
QY	61 IE-TGDARPFOQLRAQAEIAIDEADVITTFMVNVREGLTQSDENVAQILYSKKRPFVLAVN 119
DB	61 IEPENNNDIILTQMRRQAQIAEMSDVITTFMVDKGSLTDINDVAIMLRKSKRIYLAVN 120
QY	120 KYDMMERTDVDEYSLGFGEPPYRISGSHGIGLDLLDAVVSHFGEEEDPYDEDTRL 179
DB	121 KIKKNVENNNIEREYNLGIDGPVISISSQGIGIDMDDEVYNNFKSGEDEDEBEYIKIA 180
QY	180 IIGRPNVGSSLVNALIGEDRVYISNAGTTTRAIDREYSTDGODVYILDAGKRRKGKV 239
DB	181 FVCGPNNGKSSLTRNLIGEEERVYISDPJGTTRDAIDFLETFDFGLKLYLIDTAGLRKRSRI 240
QY	240 YESREKSVLRALKAIERSNVVLVIDAEGSIIEODKRVGAHEOGKAAYIVNKMDTV 299
DB	241 KEELERISAVRTMAIERCDVCTILIDATFERISODEDEIIIGYAHENNKAILIYVNKDLI 300
QY	300 EKDSKTAKKEDEVERKEFFLDYAOIAFVASKERTRLTLFPYINEASENHKKRVOSTL 359
DB	301 EKDKDTMENKKNLNEMKFSEMAVAPRLFISAKTGORVHKVLSLKCYDYNSKRIATGV 360
QY	360 NEVVTDAISNPPFTDGRRLNFYAQOVALEPRTFYEVVNDVELMHFSKYRYLENORA 419
DB	361 NDVISNAVLMEKPVPVAFKRLKIETGYQTDIKPTTFEFVANNPELLHFSTRYLENKLRO 420
QY	420 AFEGEGTPIHIIAKKR 435
DB	421 SFDPEGTGIKMIFFER 436
RESULT 13	
ENGAL_FUSNN	STANDARD; PRT; 440 AA.
AC	QBRGV7;
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Probable GTP-binding protein engA.
 GN ENG A OR FN0170.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 CC Fusobacterium.
 OX NCBI_TaxId=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapralov V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhatnagar A., Barman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyriades N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENG A SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE010530; AAL94376.1; -
 CC HAMAP: MF_00195; -; 1
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR TIGRfams: TIGR00650; MG442; 2.
 DR TIGRfams: TIGR00231; small_GTP; 2.
 KM GTP-binding; Repeat; Complete proteome.
 FT NE_BIND 9 16 GTP 1 (POTENTIAL).
 FT NE_BIND 56 60 GTP 1 (POTENTIAL).
 FT NE_BIND 119 122 GTP 1 (POTENTIAL).
 FT NE_BIND 183 190 GTP 2 (POTENTIAL).
 FT NE_BIND 230 234 GTP 2 (POTENTIAL).
 FT NE_BIND 295 298 GTP 2 (POTENTIAL).
 SQ SEQUENCE 440 AA: 49608 MW: BEC805F839BE2A05 CRC64;
 Query Match 53.3%; Score 1178.5; DB 1; Length 440;
 Best Local Similarity 53.6%; Pred. No. 43e-65;
 Matches 221; Conservative 81; Mismatches 116; Indels 3; Gaps 3;
 QY 3 KPIVAIVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRIRYSSGEMLTDFNITDGGIE 62
 DB 2 KPIIAIVGRPNVSKSTIFNLVGDKIAIVDLPGVTRDRIRYRTDMSGSEFVIVDGGIE 61
 QY 63 IGDAPF-QTQIRAOAEIAIDEADVIIFMNVNREGILQSDMAVQILKSKRPVLAIVNKV 121
 DB 62 PRNNDLMAIKIKQAEVAMNEADVIIFVVDGKSGMLPDEIAYIIKRNKRPVILCVNKI 121
 QY 122 DN-MENRTVDYDYPSLGFGEPIYSGSHGLGLDLDAVYSHRGEEDPDYDTRLSI 180
 DB 122 DNFEEQDDYDYDYGIGFELVPISGHKYVLDMDLDIYVDIIGKMDPFEDEVLKLV 181
 QY 181 IGRPNVKSLLVNAIIGEDRIVSVNAGTTRDAIDTEYSYDQDYVLDITAGARRKGVY 240
 DB 182 IGRPNVKSLLVNAIIGEDRIVSVNAGTTRDAIDTEYSYDQDYVLDITAGARRKGVY 241
 QY 241 ESTEKSIVLRALKATRSNVVIVDAEQCIIEQDKRVACVYAHQGNVIVVNMKDTVE 300
 DB 242 ESTEYSVLRALKATRSNVVIVDAEQCIIEQDKRVACVYAHQGNVIVVNMKDTVE 301
 QY 301 -KDSKMKKFEEDVRRKFEFDVDAQIAFVSAKERTLTLTFFIINASEMHHKRVOSSTL 359
 DB 301 -KDSKMKKFEEDVRRKFEFDVDAQIAFVSAKERTLTLTFFIINASEMHHKRVOSSTL 359

DB 302 NKNNAIKKIKKEELVELPLSLAPIEFVSALIGQRTNLEADRIYEETKRISTGL 361
 QY 360 NEVVTDASMPPTDGRRLNVEYATQVAIEPPTFVFNVDLMHFYSKRYLENOIRA 419
 DB 362 NTLIKDVLMMNPPTRGRIKINAYQVAVAPKPFVFCNYPILHFSARYIENKRE 421
 QY 420 AFGEGTPIHI 430
 DB 422 AFGEGSPIMI 432
 RESULT 14
 ENG_A_NASP STANDARD; PRT; 453 AA.
 ID ENG_A_NASP
 AC 08YZH7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Probable GTP-binding protein engA.
 GN ENG A OR ALR0483.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxId=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENG A SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP003582; BAB72441.1; -
 CC PIR: AB1867; AB1867.
 DR HAMAP: MF_00195; -; 1.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP1_OBG.
 DR InterPro: IPR002917; MMR_HSR1.
 DR InterPro: IPR001806; Ras_transf.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF01926; MMR_HSR1; 1.
 DR PRINTS: PR00326; GTP_OBG.
 DR PRINTS: PR00449; RASTRNSFRMG.
 DR TIGRfams: TIGR00650; MG442; 2.
 DR TIGRfams: TIGR00231; small_GTP; 2.
 KM GTP-binding; Repeat; Complete proteome.
 FT NE_BIND 10 17 GTP 1 (POTENTIAL).
 FT NE_BIND 57 61 GTP 1 (POTENTIAL).
 FT NE_BIND 120 123 GTP 1 (POTENTIAL).
 FT NE_BIND 183 190 GTP 2 (POTENTIAL).
 FT NE_BIND 230 234 GTP 2 (POTENTIAL).
 FT NE_BIND 295 298 GTP 2 (POTENTIAL).
 SQ SEQUENCE 453 AA: 50731 MW: 23432F10C2E287D1 CRC64;
 Query Match 46.9%; Score 1037; DB 1; Length 453;
 Best Local Similarity 47.6%; Pred. No. 2e-56;
 Matches 210; Conservative 83; Mismatches 138; Indels 10; Gaps 5;
 QY 1 MKPIVAIVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRIRYSSGEMLTDFNITDGG 60
 DB 1 MKPIVAIVGRPNVSKSTIFNRIVGERVSIYEDTPGVTRDRIRYSSGEMLTDFNITDGG 60

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 00:29:31 ; Search time 100 Seconds
(without alignments)
1125.109 Million cell updates/sec

Title: US-09-815-242-12600
Perfect score: 2213
Sequence: 1 MTKPIVAIVGRPNVNGKSTIF.....IRAAFGEGTPIHIIAKRN 436

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirts:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	2094	94.6	436 16 O8CP62	O8CP62 staphylococ
2	1614	72.9	436 16 O8EQA8	O8EQA8 oceanobacil
3	1523	68.8	436 16 O8E3T9	O8E3T9 streptococ
4	1523	68.8	436 16 O8DY73	O8DY73 streptococ
5	1476	66.7	436 2 O9RHV5	O9RHV5 streptococ
6	1472	66.5	436 16 O8DS90	O8DS90 streptococ
7	1342	60.6	445 2 O8CE58	O8CE58 heliobacill
8	1340.5	60.6	435 2 O8KH12	O8KH12 lactobacill
9	1028.5	46.5	449 16 O8DK11	O8DK11 synechococ
10	1022	46.2	444 16 O8EWH6	O8EWH6 mycoplasma
11	920	41.6	503 16 O8FF59	O8FF59 escherichia
12	898.5	40.6	496 16 O8DF02	O8DF02 vibrio vuln
13	874	39.5	537 10 O9LH58	O9LH58 arabidopsis
14	874	39.5	659 10 O9CF70	O9CF70 arabidopsis
15	859	38.8	487 16 O8EC36	O8EC36 shewanella
16	851.5	38.5	463 16 O8G6A8	O8G6A8 bifidobacte

17	843.5	38.1	552 16 O8FTK5	O8FTK5 corynebacte
18	838	37.9	588 16 O9A7R6	O9A7R6 caulobacter
19	796.5	36.0	483 16 O8G2E8	O8G2E8 brucella su
20	756.5	34.2	874 5 O8I5N5	O8I5N5 plasmodium
21	736	33.3	496 10 O9AW74	O9AW74 guillardia
22	678	30.6	489 16 O8F6K1	O8F6K1 leptospira
23	661	29.9	188 11 O92212	O92212 rattus norv
24	559.5	25.3	383 2 P94645	P94645 campylobact
25	524	23.7	642 10 O9AX00	O9AX00 oryza sativ
26	514.5	23.2	514 10 O9FLE0	O9FLE0 arabidopsis
27	512	23.1	347 2 O8VN97	O8VN97 heliobacte
28	512	23.1	347 2 O8VN36	O8VN36 heliobacte
29	511	23.1	347 2 O8VN38	O8VN38 heliobacte
30	455	20.6	281 2 O9LA73	O9LA73 thioacellu
31	435	19.7	456 16 O8DIY0	O8DIY0 walgieswort
32	433.5	19.6	208 2 P72548	P72548 synechococ
33	320.5	14.5	190 2 O9AE08	O9AE08 campylobact
34	320.5	14.5	190 2 O9Q016	O9Q016 campylobact
35	319.5	14.4	190 2 O9AE06	O9AE06 campylobact
36	318.5	14.4	190 2 O9AEJ9	O9AEJ9 campylobact
37	316.5	14.3	190 2 O9AEJ7	O9AEJ7 campylobact
38	266	12.0	479 16 O8DF28	O8DF28 streptococ
39	258	11.7	300 16 O8EPY0	O8EPY0 oceanobacil
40	243.5	11.0	455 16 O8DFT8	O8DFT8 streptococ
41	241.5	10.9	170 2 O9S3C9	O9S3C9 heliobacte
42	241.5	10.9	170 2 O9S3C8	O9S3C8 heliobacte
43	241.5	10.9	170 2 O9S3C7	O9S3C7 heliobacte
44	241.5	10.9	170 2 O8VNA0	O8VNA0 heliobacte
45	241.5	10.9	170 2 O9R3H9	O9R3H9 heliobacte

ALIGNMENTS

RESULT 1	ID	PRELIMINARY:	PRT:	436 AA.
O8CP62	O8CP62			
AC	O8CP62:			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	GRP binding protein.			
GN	SEI163.			
OS	Staphylococcus epidermidis.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1282;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 12228;			
RA	Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,			
RA	Chen Z., Wen Y.,			
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AE016747; AAC04760.1; -			
KW	Complete proteome.			
SEQ	SEQUENCE 436 AA; 49024 MW; 77C74B95641D3F22 CRC64;			

Query Match	Best Local Similarity	94.6%:	Score 2094;	DB 16;	Length 436;
Matches	406;	Conservative	18;	Mismatches	12;
				Indels	0;
				Gaps	0;
OY	1	MTKPIVAIVGRPNVNGKSTIFPNRIGERSYIEDPFGVTRDRYISGSEWLTDFDNITDGG	60		
DB	1	MTKPIVAIVGRPNVNGKSTIFPNRIGERSYIEDPFGVTRDRYISGSEWLTDFDNITDGG	60		
OY	61	IEIGDAPFOTOIRAOAETIAIDADVITFNVNREGLTQSDENVAQILYKSKRPVLAANK	120		
DB	61	IEIGDAPFOTOIRAOAETIAIDADVITFNVNREGLTQSDENVAQILYKSKRPVLAANK	120		
OY	121	VDNMEMRTDYDFSLGCEGEPYPSGSHGLGLDLDVAVSHFGEEDDPDEDITRISI	180		
DB	121	VDNMEMRTDYDFSLGCEGEPYPSGSHGLGLDLDVAVSHFGEEDDPDEDITRISI	180		
OY	181	IGRPVNGKSLVNMILGEDRVIVSNVACTTDAIDTETYSYDGDYVLIDTAGMKRKRVY	240		

```
Db 181 IGRPNVAGSSLVNALLGGERVYVSNVAGTTTADITETSYDQDYLIDTAGMRKKGY 240
Oy 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
Db 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
Oy 301 KDSKTMKRFEDVEVRKEFOFLDYAOIAFYSAKERTLRILFPYINAESENHKKRVOSTLN 360
Db 301 KDSKTMKRFEDVEVRKEFOFLDYAOIAFYSAKERTLRILFPYINAESENHKKRVOSTLN 360
Oy 361 EVVTDIAISMNPPTDKGRRLNVEFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
Db 361 EVVTDIAISMNPPTDKGRRLNVEFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
Oy 421 FGEGCTPIHIIARKRN 436
Db 421 FGEGCTPIHIIARKRN 436
```

RESULT 2

```
O8EQA8 PRELIMINARY: PRT; 436 AA.
ID O8EQA8:
AC O8EQA8:
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE GTP binding protein.
GN O8197.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_Taxid=182710.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL: AP004599; BAC13753.1; -.
KW Complete proteome.
SQ SEQUENCE 436 AA: 48998 MW: 18P52E40D9B0C9C2 CRC64:
```

Query Match 72.9%; Score 1614; DB 16; Length 436;

Best Local Similarity 69.4%; Pred. No. 5,8e-94;

Matches 302; Conservative 66; Mismatches 67; Indels 0; Gaps 0;

```
Oy 1 MTKPIVAIVGRPNVCKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEMLTDFNIIIDGG 60
Db 1 MKKSVALVGRPNVCKSTIFNRLVGERISIVEDIQVTRDRIRYAAEMLNHFNIIIDGG 60
Oy 61 IEIGDAPPTQIIRAQAEIAIDEADYIIFMVNVRREGTLQSDENVAOILYKSKRPVLAANK 120
Db 61 IEIGDAPPTQIIRAQAEIAIDEADYIIFMVNVRREGTLQSDENVAOILYKSKRPVLAANK 120
Oy 121 VDNMEMRTDVFYFSLGGEPRPIISGSHGLGLDLDAVSHFGEEEDPYDEDTIRLSI 180
Db 121 VDNMEMRTDVFYFSLGGEPRPIISGSHGLGLDLDAVSHFGEEEDPYDEDTIRLSI 180
Oy 181 IGRPNVAGSSLVNALLGGERVYVSNVAGTTTADITETSYDQDYLIDTAGMRKKGY 240
Db 181 IGRPNVAGSSLVNALLGGERVYVSNVAGTTTADITETSYDQDYLIDTAGMRKKGY 240
Oy 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
Db 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
Oy 301 KDSKTMKRFEDVEVRKEFOFLDYAOIAFYSAKERTLRILFPYINAESENHKKRVOSTLN 360
Db 301 KDSKTMKRFEDVEVRKEFOFLDYAOIAFYSAKERTLRILFPYINAESENHKKRVOSTLN 360
```

```
Oy 361 EVVTDIAISMNPPTDKGRRLNVEFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
Db 361 EVVTDIAISMNPPTDKGRRLNVEFYATQVAIEPPTFVFVNDVELMHFSYKRYLENOIRAA 420
Oy 421 FGEGCTPIHIIARKRN 435
Db 421 FGEGCTPIHIIARKRN 435
```

RESULT 3

```
O8E379 PRELIMINARY: PRT; 436 AA.
ID O8E379:
AC O8E379:
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN GBS1667.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangoul L.,
RA Msadek T., Zouine M., Couve E., Lallouf L., Poyart C., Tiliou-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL: AL766852; CAD47326.1; -.
RA Sagalistic; gbs1667; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 436 AA: 48880 MW: E7E33B4D0F3F61Q*GRC64:
```

Query Match 68.8%; Score 1523; DB 16; Length 436;

Best Local Similarity 66.5%; Pred. No. 3.2e-88;

Matches 290; Conservative 70; Mismatches 74; Indels 2; Gaps 2;

```
Oy 1 MTKPIVAIVGRPNVCKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEMLTDFNIIIDGG 60
Db 1 MTKPIVAIVGRPNVCKSTIFNRIYGERVSIYEDTPGVTRDRIRYSGEMLTDFNIIIDGG 60
Oy 61 IEIGDAPPTQIIRAQAEIAIDEADYIIFMVNVRREGTLQSDENVAOILYKSKRPVLAANK 120
Db 61 IEIGDAPPTQIIRAQAEIAIDEADYIIFMVNVRREGTLQSDENVAOILYKSKRPVLAANK 120
Oy 121 VDNMEMRTDVFYFSLGGEPRPIISGSHGLGLDLDAVSHFGEEEDPYDEDTIRLSI 180
Db 121 VDNMEMRTDVFYFSLGGEPRPIISGSHGLGLDLDAVSHFGEEEDPYDEDTIRLSI 180
Oy 181 IGRPNVAGSSLVNALLGGERVYVSNVAGTTTADITETSYDQDYLIDTAGMRKKGY 240
Db 181 IGRPNVAGSSLVNALLGGERVYVSNVAGTTTADITETSYDQDYLIDTAGMRKKGY 240
Oy 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
Db 241 ESTEKYSVLRAKALAIERSNVVLVIDAEGGIIEDOKRVAGYAHEOGKAVVYNKMDTVE 300
Oy 301 KDSKTMKRFEDVEVRKEFOFLDYAOIAFYSAKERTLRILFPYINAESENHKKRVOSTLN 360
Db 301 KDSKTMKRFEDVEVRKEFOFLDYAOIAFYSAKERTLRILFPYINAESENHKKRVOSTLN 360
```


RESULT 4	ID	PRELIMINARY:	PRJ:	436 AA.
Q8DY73	O8DY73			
AC	O8DY73;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Phosphoglycerate dehydrogenase-related protein.			
GN	SAG1620			
OS	Streptococcus agalactiae (serotype V).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=216466;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=2603 V/R / Serotype V;			
RX	MEDLINE=22222988; PubMed=12200547;			
RA	Tetelijn H., Masiagniti V., Ciesliewicz M.J., Eisen J.A., Peterson S.,			
RA	Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,			
RA	Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,			
RA	Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,			
RA	Radune D., Fedorova N.B., Scanlan D., Khouli H., Mulligan S.,			
RA	Carty H.A., Cline R.T., Van Aken S.E., Gill J., Searelli M., Mora M.,			
RA	Iacobini E.T., Brettoni C., Gelli G., Mariani M., Vegli F., Malone D.,			
RA	Rinaudo D., Rappelli R., Telford J.L., Kasper D.L., Grandi G.,			
RA	Fraser C.M.;			
RT	"Complete genome sequence and comparative genomic analysis of an			
RT	emerging human pathogen, serotype V <i>Streptococcus agalactiae</i> ."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).			
DR	EMBL: AE014265; AAN00484.1; -			
DR	TIGR: SAG1620; -			
KM	Complete proteome.			
Q8	SEQUENCE 436 AA; 49881 MW; E2C064B3F04B8644 CRC64;			

Query Match:	68.8%	Score 1523:	DB 16,	Length 436;
Best Local Similarity	66.3%	Pred. No. 3.2e-88;		
Matches 289;	Conservative 73;	Mismatches 73;	Indels 2;	Gaps 2;

[illegible]

ID	09RHV5	PRELIMINARY:	PRT:	436 AA.
AC	09RHV5.			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	Phosphoglycerate dehydrogenase.			
GN	pGda.			
OS	Streptococcus mutans.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
CC	Streptococcus			
OX	NCBI_TaxID=1309;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MT8148.			
RA	Kawabata S., Terao Y., Hamada S.;			
RT	"Molecular cloning, sequence and characterization of a novel			
RT	streptococcal phosphoglycerate dehydrogenase gene.";			
RL	Oral Microbiol. Immunol. 15:58-62(2001).			
DR	EMBL, AB016077, BAA88823.1; -			
DR	InterPro: IPR005289; GTP-binding_dom.			
DR	InterPro: IPR006073; GTP1_OBG.			
DR	InterPro: IPR002917; MMR_HSR1.			
DR	InterPro: IPR005225; Small_GTP.			
DR	Pfam: PFO1926; MMR_HSR1. 1.			
DR	PRINTS: PR00326; GTP_OBG.			
DR	TIGRFAMS: TIGR00650; MG442; 2.			
DR	TIGRFAMS: TIGR00231; small_GTP; 2.			
SO	SEQUENCE 436 AA; 48601 MW; EC52CC8650DED090 CRC64;			

Query Match	66.7%;	Score 1476;	DB 2;	Length 436;
Best Local Similarity	62.8%;	Pred. No. 3e-85;		
Matches 274;	Conservative 83;	Mismatches 77;	Indels 2;	Gaps 2

```

OY      1  MTKPIVAIVGRPNVSKSTIFNRIVERVISEDPTGVRDRIRYSSEGMWLTHTDPIITDGG  60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  MALPPIVAIVGRPNVSKSALFNRINAGERISIVEDVGCYTRDRIRYTAEMLNRPFSIIDTGG  60

OY      61  IEIGDAPFOTOIRAOAETIAIDEADVIIFMVNVREGILTOSDEMVAOILKSKKRPVYLAIVK  120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  IDDVAPRPMEOQKHQADIAIAMEADVIIVFYSAKEGITDADEVAKIILYRTHKRPVILAIVK  120

OY      121  VDNEMERTDYIDFYSLGFEERYPISGSHGLGILLDAVYSHFGEEDBDPYEDPIIRLSI  180
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121  VDNPEMRSAIYDFVALGLGDPYPVSSAHGIGTGVDLAIVDMLPADEGE-SSDIKKFSL  179

OY      181  IGRPNVGSLLVNALIGEDRVIVSNVAGTTRDAIDETS-XUGODYULDTGMRKKKGV  239
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      180  IGRPNVGSLLNALIGEDRVIVASPVAGTTRAIDTTFTEDEGQETTMIDTGMRSKRGV  239

OY      240  YESTEKISYVLRALKAIERSNVVLYVYIDAEOGLIEODKRAGAGAHDEGAKVNVLYVKNMPTV  299
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      240  YENTEKISYVMRMRADIDRSSDIYALMLNAEEGIREYDKRIRAGEAHEGAKGIVVYVKNMAI  299

OY      300  EKDKSTIMKKFEDEVAKEFOFLDYADAIAPYSAEKERLRRLTFYIINDEASNHHKXROSSVL  359
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      300  KKDNRIVAQWETDIDNFOYIPYAIIVFYSAATKORLKRLPVIKOVSSQOMTRIRPSSVL  359

OY      360  NEVYVDAISMNPTPIIDKGRILVAFATOAIEPRPFVYVNVDELIMHFSYKRYLENOJRA  419
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      360  NDYVDAVAIINTPIIDKGRILKIFATOVSVKPRPFVIVNVEEELMHFSYLFLENOJHO  419

```

QY	420	AFEGEGTPIHIAARR	435
		::	
Db	420	AFVEEGTPIHIAARR	435

RESULT 6	
Q8DS90	
ID Q8DS90	PRELIMINARY;
AC Q8DS90;	PRT; 436 AA.
DT 01-MAR-2003	(TREMBLrel. 23, Created)
DT 01-MAR-2003	(TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Phosphoglycerate dehydrogenase.
 GN PGD OR SMU.1920.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN NCBI_TaxID=1309;
 RX NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN-UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL: AE015016; MAM59531.1; -.
 KW Complete proteome.
 SQ SEQUENCE 436 AA; 48585 MW; 24D8428A91C2A097 CRC64;

Query Match 66.5%; Score 1472; DB 16; Length 436;
 Best Local Similarity 62.6%; Pred. No. 5.3e-85;
 Matches 273; Conservative 84; Mismatches 77; Indels 2; Gaps 2;

QY 1 MTKPIVAIVGRPNVSKSTIFNRIVGERSIVEDTGTGTRDRIRYSSEGLTHDFNIIDTGG 60
 DB 1 MALPVAIVGRPNVSKSALEFNRIVGERSIVEDVGTDRIRYTAEMLRQFSIIDTGG 60
 QY 61 IEIGAPFQTOIRAOAETIAIDADYITFVNVRREGITQSDENVAQILKSKRPVLA 120
 DB 61 IDDVAPFMEQIKHQADIAIMTADYIVFVVASKEGTTDADEVAKILYKTHKPVILAVNK 120
 QY 121 VDNMEKRTDYDFYSLGEGEPYPSGSHGLGLDLDAVSHFGEEDPYDEDTIRLSI 180
 DB 121 VDNPEKRSIVFYALGLGDPFYPVSSAHGIGTGDVDAIVDNLPRDAQEE-SSDIKESL 179
 QY 181 IGRPNVGSLLVNAIIGEDRIVTSNVACTRDADITEXS-YDGOYVILDTAGMRKKGV 239
 DB 180 IGRPNVGSLLVNAIIGEDRIVTSNVACTRDADITEXS-YDGOYVILDTAGMRKKGV 239
 QY 240 YESTKYSYLRALKAIERSNVLYVIDAEOGIIEDDKRAGVAHOGKAVIVVVKMDIV 299
 DB 240 YENTKYSYLRALKAIERSNVLYVIDAEOGIIEDDKRAGVAHOGKAVIVVVKMDIV 299
 QY 300 EKDSKTKMKFEDEVKKEFOFLDYAOIAFVSAKERTRLTLPYINEASENHKKRVOSTL 359
 DB 300 KKDNTVAQMEADIDNDNQIYIPAYIVFSAVTKORLHKLPVIVKQVSSQMTIRPSAVL 359
 QY 360 NEVTDAISMNPPTDKGRRLNVFYATOVAIEPPTFVFNVDVLMHFSYKRYLENQIRA 419
 DB 360 NQVMDAVAINPTPDCKGRKRIFYATQVSVKPRPFVIVNEEELMHFSYKRYLENQIRQ 419
 QY 420 ARGFEGTPIHIIARRK 435
 DB 420 ARGFEGTPIHIIARRK 435

RESULT 7
 O8GES8 PRELIMINARY: PRT; 445 AA.
 AC O8GES8;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE GTP-binding protein (Fragment).
 OS Helicobacterium mobilis.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;
 OC Helicobacter.
 RN NCBI_TaxID=28064;
 RX SEQUENCE FROM N.A.
 MEDLINE=22337798; PubMed=12446909;

RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
 RA Blankenship R.E.;
 RT "Whole-genome analysis of photosynthetic prokaryotes";
 RL Science 298:1616-1620(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Liolios K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,
 RA Gerdes S., Kyriakides N., Overbeek R.;
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY142760; MAM87364.1; -.
 FT NON-TER 445 445
 SQ SEQUENCE 445 AA; 49885 MW; 1D6A692E387AD61C CRC64;

Query Match 60.6%; Score 1342; DB 2; Length 445;
 Best Local Similarity 56.8%; Pred. No. 8.8e-77;
 Matches 250; Conservative 83; Mismatches 99; Indels 8; Gaps 3;

QY 1 MTKPIVAIVGRPNVSKSTIFNRIVGERSIVEDTGTGTRDRIRYSSEGLTHDFNIIDTGG 60
 DB 4 MAKPIVAIVGRPNVSKSTIFNRIVGERSIVEDTGTGTRDRIRYRNKAKWLNREFTLVDTGG 63
 QY 61 IEIG--DAPFQTOIRAOAETIAIDADYITFVNVRREGITQSDENVAQILKSKRPVLA 118
 DB 64 IERGADGNPFSEYITKQAPATAIADYILFMYDGKAGITADETYAAVLRITKRYFLV 123
 QY 119 NKVDNMEKRTDYDFYSLGEGEPYPSGSHGLGLDLDAVSHF--GEEDPYDEDT 175
 DB 124 NKIEDPSQDKYEFEPFALGLGDPFISAHGNMTGOLDVAIEALPPEGDEDDP--DT 180
 QY 176 IRLSIIIGRPNVGSLLVNAIIGEDRIVTSNVACTRDADITEXSIDGQDYVILDTAGMRK 235
 DB 181 IKIAVIGRPNVGSLLVNAIIGEDRIVTSNVACTRDADITEXSIDGQDYVILDTAGMRK 240
 QY 236 KGVYESTKYSYLRALKAIERSNVLYVIDAEOGIIEDDKRAGVAHOGKAVIVVVKMDIV 295
 DB 241 KKRIDESVRYVMSKSLRAVDSVYLAVIDASQGTEDDKRIAGAHAGACVVLNKK 300
 QY 296 WDTVEKSKTKMKFEDEVKKEFOFLDYAOIAFVSAKERTRLTLPYINEASENHKKRVQ 355
 DB 301 WDLVPEKDKTKMRFQDVVSELGFLNYPATIVYSALTGRLPKILELVDFVFEQANRRIP 360
 QY 356 SSTLNEVTVDAISMNPPTDKGRRLNVFYATOVAIEPPTFVFNVDVLMHFSYKRYLEN 415
 DB 361 TSVINELMDIIRVPAPSDRGRKILVYTOTSVKPRPFVFNDEELHFSYKRYLEN 420
 QY 416 QIRAAFGFEGTPIHIIARRK 435
 DB 421 RRETFEGFEGTPIHIIARRK 440

RESULT 8
 O8KH12 PRELIMINARY: PRT; 435 AA.
 AC O8KH12;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Conserved hypothetical GTP-binding protein.
 OS Lactobacillus delbrueckii (subsp. bulgaricus).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 RN NCBI_TaxID=1585;
 RX SEQUENCE FROM N.A.
 RC STRAIN-ATCC 11842;
 RA Serrif P., Deryn R., Ehrlich S.D., Maguin E.;
 RT "Lactobacillus delbrueckii spp. bulgaricus hba region";
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY094626; MAM22484.1; -.
 DR InterPro: IPR000795; EF_GTPbind.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR006073; GTP_OBG.
 DR InterPro: IPR002917; MMR_HSR1.

QY	125	CMRT--DYVDYFSLGGERPIGSGHGLGDLDDNAVHSGFEEDERPYDEDTIRLSIGR	183
Db	124	NKOSTIDDEHYLFREGQKPRFJISANHGIGGLDLDV--TYLKIKEDDEPFCCIGR	180
QY	184	PWVGKSSLVMAIIAGEDRYIVSNVAGTTBDAIDTEYSYDQDYYLIDTAGMRKKGKYEEST	243
Db	181	PWVGKSSLVNCILINDRMITTSNIANTBDAISNEFKDMLTYIIIDTAGIRRGKIOENV	240
QY	244	EKYSYLRLLKAERESNVYLVVIDAOGGIIEDOKRVAGVHNEOGKAVIIVYNNKRDY-EKD	302
Db	241	DKYALRLRQGSOSFNLIYVLDGSEEFNEODVYAGLHKNANIPSIIVYNNKWDIYKED	300
QY	303	SKTMKKFEDEVAKKEFQDLAQIAFVSAKERTRLFTLPYINDEASNNHKRVQSYTLNEY	362
Db	301	EKTMKKFKTIRIKERKFLSWPIVEFSLAENNRIMTFLNDEIKSIRENLMLKFNKSLITDL	360
QY	363	VTDASIMNPTPTDKGRRLVFAVATOAIDPRPEVAVNVNVEMLHFSKYKRLLENQIRAAFG	422
Db	361	VFKLOMLMSPLFNNGRGRIKININVTYODGOIPFLVFCNNPEYIHFESYARILEINEIRKSLG	420
QY	423	EEGPRPIHIIARRK	435
Db	421	LNNVPITLYFKNK	433

RESULT	ID	Q8FEF59	PRELIMINARY;	PRT;	503 AA.
AC	Q8FEF59				
DT	01-MAR-2003	(TREMBLrel. 23, Created)			
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE		Probable GTP-binding protein enga.			
CN		C3033.			
OS		<i>Escherichia coli</i> O6.			
OC		Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
CC		Enterobacteriaceae; Escherichia.			
OX		NCBI_TaxID=217992;			
RP		[1]			
RP		SEQUENCE FROM N.A.			
RC		STRAIN-O6:H1 / CFT073 / ATCC 700928;			
RX		MEDLINE=22388234; Pubmed=12471157;			
RA		Weich R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,			
RA		Rasko D., Buckles E.L., Liou S.-R., Boutin A., Heckert J., Stroud D.,			
RA		Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,			
RT		Mobley H.L.T., Donnenberg M.S., Blattner F.R.;			
RT		"Extensive mosaic structure revealed by the complete genome sequence			
RL		of uropathogenic <i>Escherichia coli</i> .";			
RL		Proc. Natl. Acad. Sci. U.S.A. 99:117020-17024(2002).			
DR		EMBL; AE016764; AANB1463.1; -.			
DR		Complete proteome.			
QO		SEQUENCE 503 AA; 56573 MW; 711A9BD7863280A CRC64;			

Query Match	41.6%;	Score 920;	DB 16;	Length 503;
Best Local Similarity	42.0%;	Pred. No. 4.6e-50;		
Matches 197;	Conservative 91;	Mismatches 143;	Indels 38;	Gaps 4;

[illegible]

```

OY      216  TEYSYDQDYLVDITGMRKKGVYESTEKYSVLBALKAIERSNVLVVIDAEOGITED 275
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      256  IPMERDREYVLIDITGAVRKGKITDPAVKEFSVITTLQAIEPANVMVLVIDAREGISD 315
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY      276  KRVAGYAHBEGKAVVLVYVKKMPTVEKDSVTYMKKFEDEYRKPEOPLDVAQIAFVAKERT 335
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      316  LSLTGFLNGBRSLVLYVKKMDLSOEYKQKELLDR--LGFIDFARVHFIJALHSG 373
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY      336  LRTLPYINEASENHKKRYOQSSTLVEVYDAISNMPPTDGRRLNFEVATOVAIEPTF 395
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      374  VGNLFEEVREAYDSSTRVGTSMLTRIMTMAVEDHQPLVYGRVRKLYAHAGYNPPIV 433
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY      396  VFEVNDYELMHSEYKRYLENOIRAAFGFSGTIIH-----IARRN 436
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      434  VIHGNQVKDLPDSYKRYLNNYFRKSLDVGSGSIRLOFKEGENPYANKN 482
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 12			
ID	Q8DF02	PRELIMINARY;	PRT; 496 AA.
AC	Q8DF02;		
DT	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	GRP binding protein.		
CN	WV10423.		
OS	Vibrio vulnificus.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;		
OC	Vibrionaceae; Vibrrio.		
OX	NCBI_TaxID=672;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CMCP6;		
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.		
RT	Choy H.E.;		
RL	"Complete genome sequence of <i>Vibrio vulnificus</i> CMCP6.";		
RL	Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.		
SK	EMBL; AE016798; AA008946.1; -.		
SW	Complete proteome.		
SW	SEQUENCE 496 AA; 55469 MW; 27196BA312B7923 CRC64;		

[illegible]


```

QY 62 -----EIGDAFPQTOIRAQAEIAIDEADVIIFMVNVRG 95
DB 217 KSPSGVMEELNVSTTIGMEGIPLSRREAIAIRMPSEMEKQATAVDESAVIIFVDDQAG 276
QY 96 LTQSDENVAAQL--YKSKRRVVLAVNKVDN-MEMRTDVPDYFSLGFEPEPPIGSHGL 152
DB 277 PSADVEIAIDWLKKYSHKYLIIIAVNCESPRKGLMQASEFWSLGF--TPPIISALSGTG 335
QY 153 GDLLDAVVS-----HFGEEEDPDYEDTIRLSITGRPNVGSLSLVAILGEDRVIV 203
DB 336 GELLDLVCSGLIKLEIMENEIEEEENY---IPALITGRPNVGSLSILNALVEDRTIV 392
QY 204 SNAAGTTRDAIDTYS-YDGQDYVLIDTAGMRKKGVES--TEKYSVLRAKAIERSN 259
DB 393 SPVSGTTRDAIDAEFTGPDEKFRLLDTAGIRKSSVASSGSTTEAASVRAFAIRSD 452
QY 260 VLVIVDAEBCIIEODKRVAGVAHEOGKAVIIVNKKDVI-EKDSKMKKFEDEVREFO 318
DB 453 VVALVIEAMACITEODUKIAERIEREGKGLVYVNMKDTIPNKKQETAHAHEDDVREKLR 512
QY 319 FLDYAQIAFVSAKERTRLTLPYINEASENHKKRVOSTLNEVYTDALISM-NPTPYDKG 377
DB 513 SLKMAPIVSTAITGHSVDNIVAAATVQKERSRLSTALLNQVIRAAVAFKSPRRRG 572
QY 378 RLNVFATOVAIEPPFVYVNDVDELHMFYSKRYLENOIRAAFGFEGTPIHITARR 435
DB 573 KRGRVYCTQAAIRPPTFEFVNDAKLFSDTYRRYMEKQIRTAGFAGTPIRLIMRSR 630

```

RESULT 15

```

ID 08EC36 PRELIMINARY; PRT; 487 AA.
AC 08EC36;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE GTP-binding protein Enga.
GN ENGA OR S03308.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxId=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MR-1;
RX MEDLINE=22297686; PubMed=1236813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Galois E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impriali M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouli H., Gill J., Ullrich T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.,
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL: AE015769; AANS6306.1; -.
DR TIGR: S03308; -.
KW Complete proteome.
SQ SEQUENCE 487 AA; 54531 MW; B75F86CA1BF9C20 CRC64;

```

Query Match 38.8%; Score 859; DB 16; Length 487;
 Best Local Similarity 39.9%; Pred. No. 3,1e-46;
 Matches 180; Conservative 98; Mismatches 147; Indels 26; Gaps 4;

```

QY 4 PIVAIGRPVNGKSTINRIVGERVSTVEDTGVTRDRIRYSSGEMLTGHDNIIDTGIEI 63
DB 3 PVALVIGRPVNGKSTINRIVGERVSTVEDTGVTRDRIRYSSGEMLTGHDNIIDTGIEI 62
QY 64 GDAFQOIRAOEIAIDEADVIIFMVNVRGGLTQSDENVAAQLIYKSKRPVVLAVNKVDN 123

```

```

DB 63 TEEGIEFKMAEQSLAIEADVIIFMTDARAGLTAADLSIAOHLRSROKTFPVVANKIDG 122
QY 124 MEMRTDVPDYSLGFEPEPPIGSHGLGLDLDVAVSHF-----GEEB---EDPY 171
DB 123 IDADSACAEFWISGLGEVYQMAAAGRGVNTMIEVALTPYAEMMGIERQEEEBEDVERQY 182
QY 172 DED-----TIRLSITGRPNVGSLSLVAILGEDRVIVSNVAGTTRDAIDTYS 219
DB 183 TEEEAEBQRLDLPKILAIIGKPNVGSLSLVAILGEDRVIVSNVAGTTRDAIDTYS 242
QY 220 YDGQDYVLIDTAGMRKKGVES--TEKYSVLRAKAIERSNVLVIVDAEBCIIEODKRV 279
DB 243 RDGREYIIDTAGYRRSKVHEVIEKFSVKTAKAEDANVLLIIDAREGVAEQDLGL 302
QY 280 GVAHEOGKAVIIVNKKDVIKSKKFEDEVREFOFLDYAQIAFVSAKERTRLRL 339
DB 303 GFALNAGRALVIAVNMKWDIGQIK--DRVKSELDRLGPIDEARIHIFISALHGTGCHL 360
QY 340 FPYINEASENHKKRVOSTLNEVYTDALISMNPTPDKGRLLNFYATQVAIEPPTFV 399
DB 361 FESIEEAYDSATRVSTSLTRIMQMSQDDHOPPLVNGRRVKKLYAHAGVNPPIVYIHG 420
QY 400 NDVELHMFYSKRYLENOIRAAFGFEGTPIHI 430
DB 421 NOVSKLPDSYKRYMANFRSLKAVYGTPIOL 451

```

Search completed: September 10, 2003, 00:35:04
 Job time : 105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:16:39 : Search time 129.915 Seconds
(without alignments)
8041.284 Million cell updates/sec

Title: us-09-815-242-1463

Perfect score: 387
Sequence: 1 gatcttcctctcttcacca.....ttcactctgaagaataaac 387

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	387	100.0	387	23	AAS48886
2	387	100.0	1311	22	AAF6461
3	387	100.0	1311	23	AAF6461
4	383.8	99.2	1305	23	AAS5497
5	383.8	99.2	1311	23	AAS51646
6	327	84.5	3621	18	AAV74669
7	295	76.2	372	23	AAV50706
8	262.2	67.8	1332	24	ABN50883

9	262.2	67.8	3269	22	AAH54708
10	234	60.5	298	23	AAS50205
11	234	60.5	298	23	AAS50723
12	205.8	53.2	319630	24	ABO67194
13	205.8	53.2	2944528	24	ABO03041
14	205.8	53.2	3011208	24	ABO69245
15	174.4	45.1	1311	24	ABK75008
16	162.6	42.0	1308	24	ABN68458
17	162.6	42.0	2365589	24	ABA90521
18	160.4	41.4	1308	24	ABN68457
19	160.4	41.4	2155561	24	ABN71527
20	153	39.5	1311	21	AA54516
21	151.4	39.1	1308	25	ABX07474
22	151.4	39.1	1311	21	AAZ91826
23	151.4	39.1	5066	19	AAZ52122
24	151.4	39.1	2162598	25	AB56454
25	146	37.7	246	23	AAS49268
26	101	26.1	960	22	AAH53683
27	89.4	23.1	1512	23	AAS52517
28	89.4	23.1	11574	22	AAS46244
29	86	22.2	157	23	AAS50894
30	84.4	21.8	1512	22	AAF94379
31	84.4	21.8	1512	24	ABK64943
32	84.4	21.8	1515	23	AAS53235
33	83.8	21.7	1473	23	AAS55968
34	80.6	20.8	640681	24	ABA92787
35	80.2	20.7	1830121	17	AAT42063
36	73.2	18.9	1370	19	AAZ96377
37	73	18.9	33140	22	AAF28536
38	69.8	18.0	1146	25	ACAO0899
39	69.8	18.0	1557	22	AAH65442
40	69.8	18.0	349980	22	AAH68528
41	69.8	18.0	349980	22	AAH68529
42	64	16.5	25360	22	AAF88314
43	64	16.5	29736	22	AAF88317
44	62	16.0	580073	18	AAT58840
45	61.4	15.9	78845	21	AAA81463

ALIGNMENTS

RESULT 1
AAS48886
ID AAS48886 standard; DNA: 387 BP.
AC AAS48886;
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #110.
XX
KW Antisense: ss: prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX

S. epidermidis gen
Staphylococcus aur
Staphylococcus aur
Listeria innocua C
Listeria monocyt
Bacillus lichenifo
Streptococcus poly
Genomic sequence o
Streptococcus poly
Streptococcus poly
ypbc gene of Stre
S. pneumoniae type
Streptococcus pneu
Streptococcus pneu
Streptococcus pneu
Staphylococcus aur
S. epidermidis ope
E. coli DNA for ce
DNA encoding novel
Haemophilus influe
DNA encoding Haemo
Haemophilus influe
Salmonella typhi D
Buchnera sp. genom
Haemophilus influe
S. pneumoniae derl
Genomic fragment #
C. glutamicum derl
C. glutamicum codin
C. glutamicum codin
S. glutamicum codin
S. pneumonia DNA fra
S. sp. DNA fra
Mycoplasma genital
N. meningitidis pa

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT;
 XX Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 XX
 PT New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 1: Seq ID No 1463; 511pp; English.
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence is an antisense
 CC oligonucleotide of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 387 BP; 122 A; 91 C; 53 G; 121 T; 0 other:
 Query Match 100.0%; Score 387; DB 23; Length 387;
 Best Local Similarity 100.0%; Pred. No. 1.2e-98;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTTCCTCCCTCCACCAAAATGAGAAACACGATCTTAACAGTCACCAACACTA 60
 DB 1 GATCTTCCTCCCTCCACCAAAATGAGAAACACGATCTTAACAGTCACCAACACTA 60
 OY 61 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTTAATAGAAATCATACA 120
 DB 61 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTTAATAGAAATCATACA 120
 OY 121 CGTCTGTACGATTTCCATATTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 180
 DB 121 CGTCTGTACGATTTCCATATTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 180
 OY 181 ATTTGTATATAATTTGACGACGACATTTGATGCTTGTCAATCTTCACGACGTTAA 240
 DB 181 ATTTGTATATAATTTGACGACGACATTTGATGCTTGTCAATCTTCACGACGTTAA 240
 OY 241 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTTAATTTTG 300
 DB 241 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTTAATTTTG 300
 OY 301 TTTGGAATGTGACATCACCAATTTCAATACCACTGTATCAATATTTGAATCATGTG 360
 DB 301 TTTGGAATGTGACATCACCAATTTCAATACCACTGTATCAATATTTGAATCATGTG 360
 OY 361 TTAACCATTCACCTGGAAGAAATATAATAC 387
 DB 361 TTAACCATTCACCTGGAAGAAATATAATAC 387

RESULT 2
 AAF86461/c
 ID AAF86461 standard; DNA; 1311 BP.
 XX
 AC AAF86461;
 XX
 DT 26-JUN-2001 (first entry)
 XX

DE Staphylococcus aureus yphc coding sequence.
 XX
 XX yphc; antimicrobial; cytosolic; antiulcer; microbial infection;
 KW gene therapy; vaccine; gastrointestinal carcinoma; gastric ulcer;
 KW gastritis; ds.
 XX
 OS Staphylococcus aureus.
 XX
 XX Key Location/Qualifiers
 FH 1..1311
 FT CDS /tag= a
 FT /product= "Staphylococcus aureus yphc protein"
 XX
 XX WO200123418-A1.
 XX
 XX 05-APR-2001.
 XX
 PF 19-SEP-2000; 2000WO-US25566.
 XX
 XX 28-SEP-1999; 99US-0406968.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Zalacain M, Biswas S, Burnham MKR, Sylvester D, Mcdevitt D;
 PI Mathie TB;
 DR WPI: 2001-308138/32.
 DR P-PSDB: AAB82089.
 XX
 PT Novel yphc polypeptides of Staphylococcus aureus useful for diagnosing
 PT and treating microbial infections, especially infection by
 PT Staphylococcus aureus and Helicobacter pylori -
 XX
 XX Claim 2: Page 2-3; 41pp; English.
 CC The present sequence is the gene encoding yphc polypeptide of
 CC Staphylococcus aureus. The yphc coding sequence and protein are useful
 CC for treating and diagnosing microbial infections such as infection caused
 CC by S.aureus and Helicobacter pylori. In addition, the yphc coding
 CC sequence and protein are useful for treating diseases such as
 CC H-pylori-induced cancers, e.g. gastrointestinal carcinoma, gastric
 CC ulcers, and gastritis. The present sequence was obtained from a library
 CC of clones of chromosomal DNA of S.aureus in E.coli. The sequencing data
 CC from two or more clones comprising overlapping S.aureus DNAs was used to
 CC construct the present contiguous DNA sequence.
 SO Sequence 1311 BP; 451 A; 184 C; 278 G; 398 T; 0 other:
 Query Match 100.0%; Score 387; DB 22; Length 1311;
 Best Local Similarity 100.0%; Pred. No. 1.8e-98;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATCTTCCTCCCTCCACCAAAATGAGAAACACGATCTTAACAGTCACCAACACTA 60
 DB 508 GATCTTCCTCCCTCCACCAAAATGAGAAACACGATCTTAACAGTCACCAACACTA 449
 OY 61 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTTAATAGAAATCATACA 120
 DB 448 AACCATGTGACCCGTATATCGGATACGGTTCACCAAAATCTTAATAGAAATCATACA 389
 OY 121 CGTCTGTACGATTTCCATATTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 180
 DB 388 CGTCTGTACGATTTCCATATTATCTACTTTGTTAACCCTAATACGACCGGTTTTTAA 329
 OY 181 ATTTGTATATAATTTGACGACGACATTTGATGCTTGTCAATCTTCACGACGTTAA 240
 DB 328 ATTTGTATATAATTTGACGACGACATTTGATGCTTGTCAATCTTCACGACGTTAA 269
 OY 241 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTTAATTTTG 300
 DB 268 CCATATAAATATATACATCCGCTTCATATGAGCGATTTGCGCTCTTAATTTTG 209

PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
PS Claim 1; Seq ID No 3283; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 372 BP; 113 A; 83 C; 51 G; 125 T; 0 other;
XX
Query Match 76.2%; Score 295; DB 23; Length 372;
Best Local Similarity 100.0%; Pred. No. 7e-73;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 93 CCAATACCTAATGAATGAAATCATACAGCTGTCAGCATTTCCATATTACTCTTG 152
DB 1 CCAATACCTAATGAATGAAATCATACAGCTGTCAGCATTTCCATATTACTCTTG 60
OY 153 TTAACCGCTAATACGACCGGTTTTTGTAGATTGTATAAATTTGACGACCATTTGTCG 212
DB 61 TTAACCGCTAATACGACCGGTTTTTGTAGATTGTATAAATTTGACGACCATTTGTCG 120
OY 213 CTTTGTGTAATCTCTACGACGCTTAACCATTAATAATTAATCCGCTTCATCTATG 272
DB 121 CTTTGTGTAATCTCTACGACGCTTAACCATTAATAATTAATCCGCTTCATCTATG 180
OY 273 GCGATTTCGCGCGGCTCTAATTGTGTGGATGGTGATCACCATTTCATTAACCA 332
DB 181 GCGATTTCGCGCGGCTCTAATTGTGTGGATGGTGATCACCATTTCATTAACCA 240
OY 333 CCGTATCAATATATGAAATCATGTGTTAACCATTCACCTGAAAGATTAATAC 387
DB 241 CCGTATCAATATATGAAATCATGTGTTAACCATTCACCTGAAAGATTAATAC 295
XX
RESULT 8
ABN90883/c
ID ABN90883 standard; DNA; 1332 BP.
XX
AC ABN90883;
XX
XX 24-JUL-2002 (first entry)
XX
DE *Staphylococcus epidermidis* ORF nucleic acid sequence SEQ ID NO:346.
XX
XX
KM *Staphylococcus epidermidis*: open reading frame; ORF, bacterial infection;
XX antibacterial; gene therapy; gene; ds.
OS
XX *Staphylococcus epidermidis*.
XX
PN US6380370-B1.
XX

PD 30-APR-2002.
XX
XX 13-AUG-1998; 98US-0134001.
XX
XX 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
PI WPI: 2002-381255/41.
XX
XX P-PsDB; ABP38338.
DR
XX
XX Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*
PT polypeptide, useful for diagnosing and treating bacterial infections -
PS Disclosure: SEQ ID 346; 267pp; English.
XX
XX ABN90538 to ABN93374 represent *Staphylococcus epidermidis* open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly *S. epidermidis* infections. The sequences can be used to
CC screen for compounds able to interfere with the *S. epidermidis* life
CC cycle or inhibit *S. epidermidis* infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 1332 BP; 465 A; 190 C; 267 G; 410 T; 0 other;
XX
Query Match 67.8%; Score 262.2; DB 24; Length 1332;
Best Local Similarity 79.8%; Pred. No. 1.6e-63;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
XX
OY 1 GATCTTCTCTCTCTCTGACCAAAATGAGAAACATGCTATTAACAAGTCAACCAAGCTTA 60
DB 529 GATCTTCTCTCTCTCTGATTAAGTTTCAACAACATGCTATTAACAAGTCAACCAAGTCTTA 470
OY 61 AACCATGACCGCTGATTCGATAGCGTTCCACCAANTCTAATGAATGAATTCATTA 120
DB 469 ATCCATGTGAACCCAGAAATGAGATGATGATCTCCAAAGCCTAAGAAATGAAATCATAG 410
OY 121 CGTCTGACGATTCCTCATATTAATCTACTTTGTTAACCGCTAATACGACGCTTTTGTAG 180
DB 409 TATCATTAACGATTTCAAGATTTATCAACTTATTAACAGCTATTAACAAGGTTTCTTAG 350
OY 181 ATTTGTATTAATTTGAGCGACCATTTTCATCGCTTTGTGTCAATCTTCACGACGCTTA 240
DB 349 ATTTGTATTAATTTGAGCGACCATTTTCATCGCTTTGTGTCAATCTTCATCAATTA 290
OY 241 CCATTAATAATTAATACATCCGCTTCATTCATGCGGATTTCTGCTGCGCTCAATTTGAG 300
DB 289 CCATTAATAATTAATACATCCGCTTCATTCATGCGGATTTCTGCTGCGCTCAATTTGAG 230
OY 301 TTTGGAATGCTGATCACCATAATTTCAATACCACTGTATCAATTAATTTGAATCATCTG 360
DB 229 TTTGGAATGAGAGATCTCCATATTTCAATTAACCACTGTATCAATTAATTTGAATCATCTG 170
OY 361 TTAACCATTCACCTGAAGAAATTAATAC 387
DB 169 TTAACCATTCACCTGAAGAAATTAATAC 143
XX
RESULT 9
AAH54708
ID AAH54708 standard; DNA; 3269 BP.
XX
XX AAH54708;
XX
XX 03-SEP-2001 (first entry)
XX

XX 5. epidermidis genomic polynucleotide sequence SEQ ID NO:4072.
DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX vaccination; endocarditis; ds.
KW Staphylococcus epidermidis.
XX WO200134809-A2.
XX 17-MAY-2001.
XX 09-NOV-2000; 2000WO-US30782.
XX 09-NOV-1999; 99US-0164258.
XX (GLAXO) GLAXO GROUP LTD.
XX Kimmery WJ;
XX WPI; 2001-316495/33.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX Claim 8; Page 1757-1759; 2188pp; English.
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH52971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are present in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
XX Sequence 3269 BP; 986 A; 616 C; 451 G; 1216 T; 0 other;
SQ
Query Match 67.8%; Score 262.2; DB 22; Length 3269;
Best Local Similarity 79.8%; Pred. No. 2.1e-63;
Matches 309; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 1 GATCTTCTCTCTTCAACCAAAATGAGAAACACTGCATTAACAGTACACAGACTA 60
DB 2297 GATCTTCTGATTCCTTTAATAAGTTTCAACACACTGCATTAACAGTACACAGT 2356
QY 61 AACCATGTGACCCGATATCGGATACGTTTCAACCAATCTTAATGAATAGAAATCA 120
DB 2357 ATTCATGTGAACCGAAATAGATATGATCTCCAAACCCCTAAAGAAATCAATGA 2416
QY 121 CGTCTGATGACGATTCATATTAATCTTGTAAACCGCTAATACGCGGTTTGTAG 180
DB 2417 TATCATTTAGCATTTCAAGATTAATCACTTATTCACAGCTAATCAACAGGTTCTTAG 2476
QY 181 ATTTGTATAAATTTGACGACATTTTCATGCTTTGTGTAATCTTCAACGACGTTAA 240
DB 2477 ATTATATAAGCATTTGTCGACCATTTGCTGCTTGTGAAGTCTCTCTCAACATTTGA 2536
QY 241 CCATATAAATAATTAACATCCGCTTCATATGAGGATTTGCTGCGGCTCTAATTTTG 300
DB 2537 CCATATAAATAATGATGACATCTGCTTCTCAATTTGCTGCTGACGATTTGAG 2596

QY 301 TTGGAATGTCGATCACCATTTCATATACGACCTGTATCATATATTTGAATCATG 360
DB 2597 TTGGAAGGACGATCTTCAATTAATACACCTGTATCATATGATTTAAATCATG 2656
QY 361 TTACCATTCACCTGAGAAATTAATAC 387
DB 2657 TTACCACTCGCGAGATGAATTAATAC 2683
RESULT 10
AAS50205
ID AAS50205 standard; DNA; 298 BP.
XX
XX AAS50205;
XX
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation inhibitory sequence #1429.
DE
XX
XX Antisense; ss; prokaryotic cellular proliferation;
XX antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 1; Seq ID No 2782; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms.
XX oligonucleotide of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp://wipo.int/pub/published_pct_sequences.
SQ Sequence 298 BP; 91 A; 69 C; 43 G; 95 T; 0 other;

Query Match 60.5%; Score 234; DB 23; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.9e-56;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 60
|||||
65 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 124
|||||
61 AACCATGTGACCCCTGATATCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 120
|||||
125 AACCATGTGACCCCTGATATCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 184
|||||
121 CGTCTGTACGCAATTCATATATCTACTTTGTAAACCGCTAATACGACCGTTTGTAG 180
|||||
185 CGTCTGTACGCAATTCATATATCTACTTTGTAAACCGCTAATACGACCGTTTGTAG 244
|||||
181 ATTTGTATATAAATTTGAGGACCAATTCATCGCTTGTGTCAATCCTTCACGCA 234
|||||
245 ATTTGTATATAAATTTGAGGACCAATTCATCGCTTGTGTCAATCCTTCACGCA 298
|||||

RESULT 11
AAS50723
ID AAS50723 standard; DNA; 298 BP.
AC AAS50723;
XX
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1947.
XX
XX
KW Antisense; ss: prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
XX
PD 27-SEP-2001.
XX
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;
XX
XX
DR WPI; 2001-611495/70.
XX
XX
PS Claim 1: Seq ID No 3300; 511pp; English.
XX
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acid can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 298 BP; 91 A; 69 C; 43 G; 95 T; 0 other;

Query Match 60.5%; Score 234; DB 23; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.9e-56;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 60
|||||
65 GATCTTCCTTCCTTCACCAAAATGAGAAACACTGCATCTTAACAGTCACCAAGACCTA 124
|||||
61 AACCATGTGACCCCTGATATCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 120
|||||
125 AACCATGTGACCCCTGATATCGATACGTTTCAACCAATCTTAATGAATGAATCATYACA 184
|||||
121 CGTCTGTACGCAATTCATATATCTACTTTGTAAACCGCTAATACGACCGTTTGTAG 180
|||||
185 CGTCTGTACGCAATTCATATATCTACTTTGTAAACCGCTAATACGACCGTTTGTAG 244
|||||
181 ATTTGTATATAAATTTGAGGACCAATTCATCGCTTGTGTCAATCCTTCACGCA 234
|||||
245 ATTTGTATATAAATTTGAGGACCAATTCATCGCTTGTGTCAATCCTTCACGCA 298
|||||

RESULT 12
AB067194/C
ID AB067194 standard; DNA; 319630 BP.
XX
XX
AC AB067194;
XX
XX
DT 29-AUG-2002 (first entry)
XX
XX
DE Listeria innocua config DNA sequence #7.
XX
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KW infection; ds.
XX
XX
OS Listeria innocua.
XX
PN WO200228891-A2.
XX
XX
PD 11-APR-2002.
XX
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
XX
PI Kunst F, Glaser P;
XX
XX
DR WPI; 2002-332479/37.
XX
XX
XX
XX
XX
XX
PS Claim 5: SEQ ID 7; 180pp; French.
XX
XX
CC The present invention relates to nucleic acid sequences
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be

CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of *Listeria* (potential therapeutic agents), also for
CC treating infections by *Listeria*, and are useful as immunogens in
CC anti-*Listeria* vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at [ftp://wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

XX Sequence 319630 BP; 105207 A; 55428 C; 66726 G; 92263 T; 6 other:

Query Match 53.2%; Score 205.8; DB 24; Length 319630;

Best Local Similarity 70.9%; Pred. No. 5,5e-47;

Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

3 TCTTCTCTCTCTTCCACCAAAATGAGAACAACTGCATCTACAGTCACCAAGCCTTAA 62

51123 TCCTTCTCTCTCTTGGAAAATGACAGACAGACAGATACAGATACCAAGCCTTACT 51064

63 CCATGTGACCCCTGATATCGGATACCGCTTCCACCAAAATCTTAATGAATAGAAATCAATACAG 122

51063 CCATGAGAACCGAATTTGATACCGCTTCCACCAAAATGAGAAATGAATCAATATAC 51004

123 TCTGTACCATTTTCCATATATCTACTTTGTTAACCGCTTAATACAGCGGTTTTTACAT 182

51003 TGAATCAGCATTTTGGGGTTATCTACTTTATTAATCGCTAAACAAATGGTTTATAGAC 50944

183 TTGTATTAATTAATACATCGCTTCAATCTATGGGATTTTGCCTGGCGCTTAATTTGTGT 242

50943 CGGTAAAGAAATTTTGTCTACTTGTCTGCTGATCGGTACCGCTTACAGCAATTTGGTA 50884

243 ATAAATTAATTAATACATCGCTTCAATCTATGGGATTTTGCCTGGCGCTTAATTTGTGT 302

50883 ATAAATTAATTAATACATCGCTTCAATCTATGGGATTTTGCCTGGCGGAAATTTGCTCT 50824

303 TGGATGTGTGATCAATCAATTTCAATACCACTGTATCAATATTAATGAATCAATGTGT 362

50823 AAGATGTGTGTGCTGGAAGATCAATACCACTGTATCAATATTAATTTTATTTGCA 50764

363 AACCATTCACCTGGAAGATTAATAC 387

50763 AGCCATTCGCTGGAATTAATATGCG 50739

RESULT 13

ABA03041 ID ABA03041 standard; DNA: 2944528 BP.

ABA03041; ABA03041;

05-FEB-2002 (first entry)

Listeria monocytogenes EGD-e genome sequence.

Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

vitamin B12; bacterial infection; disease; ds.

Listeria monocytogenes.

WO200177335-A2.

18-OCT-2001.

11-APR-2001; 2001WO-FR01118.

11-APR-2000; 2000FR-0004629.

(INSP) INST PASTEUR.

Buchrieser C, Frangeul L, Couve E, Rusniok C, Eshti H, Dehoux P;

Dusurget O, Chetoui F, Nedjati H, Glaeser P, Kunst F, Cossart P;

Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

Domínguez-Bernal G, Garrido-García P, Tlerréz-Martínez A, Amend A;

Chakraborty T, Domann E, Hain T, Berche P, Chardit A, Durant L;

PI Perez-Diaz J, Baguero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;

DR WPI; 2002-010914/01.

XX Genomic sequence for *Listeria* monocytogenes, useful e.g. for treatment

XX and prevention of *Listeria* and related bacterial infections, and

XX related polypeptides

PS Claim 1; SEQ ID No 1; 192pp; French.

XX The present sequence is the genome sequence of *Listeria* monocytogenes

XX EGD-e. This sequence and fragments of this sequence are useful for

XX selecting probes and primers for detecting genes in *L. monocytogenes* and

XX related organisms, and to study genetic polymorphisms and other genomes.

CC Proteins (AB047297-AB050149) expressed from the present sequence are

CC useful for raising specific antibodies, identification of *L.*

CC monocytogenes and related organisms, and for biosynthesis and

CC biodegradation, especially biosynthesis of Vitamin B12. This sequence and

CC proteins encoded by it are also useful for selecting compounds that

CC regulate gene expression and cell replication and modulate *L.*

CC monocytogenes-related diseases. In addition, this sequence and proteins

CC encoded by it are useful in pharmaceutical and vaccines compositions for

CC the treatment or prevention of infections by *L. monocytogenes* and related

CC organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at [ftp://wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

XX Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other:

Query Match 53.2%; Score 205.8; DB 24; Length 2944528;

Best Local Similarity 70.9%; Pred. No. 1.1e-46;

Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

3 TCTTCTCTCTCTTCCACCAAAATGAGAACAACTGCATCTACAGTCACCAAGCCTTAA 62

2011933 TCCTTCTCTCTCTTGGAAAATGACAGACAGATACAGATACCAAGCCTTACT 2011992

63 CCATGTGACCCCTGATATCGGATACCGCTTCCACCAAAATCTTAATGAATAGAAATCAATACAG 122

2011993 CCATGGAACCGAATTTGATACCGCTTCCACCAAAATGAGAAATGAATCAATATAC 2012052

123 TCTGTACCATTTTCCATATATCTACTTTGTTAACCGCTTAATACAGCGGTTTTTACAT 182

2012053 TGAATCAGCATTTTGGGGTTATCTACTTTATTAATCGCTAAACAAATGGTTTATAGAC 2012112

183 TTGTATTAATTAATACATCGCTTCAATCTATGGGATTTTGCCTGGCGCTTAATTTGTGT 242

2012113 CGGTAAAGAAATTTTGTCTACTTGTCTGCTGATCGGTACCGCTTACAGCAATTTGGTA 2012172

243 ATAAATTAATTAATACATCGCTTCAATCTATGGGATTTTGCCTGGCGCTTAATTTGTGT 302

2012172 ATAAATTAATTAATACATCGCTTCAATCTATGGGATTTTGCCTGGCGGAAATTTGCTCT 2012232

303 TGGATGTGTGATCAATCAATTTCAATACCACTGTATCAATATTAATGAATCAATGTGT 362

2012233 AAAAATGTTCATCGGAAGATCAATACCACTGTATCAATATTAATTTTATTTGCA 2012292

363 AACCATTCACCTGGAAGATTAATAC 387

2012293 AGCCATTCGCTGGAATTAATATGCG 2012317

RESULT 14

AB069245 ID AB069245 standard; DNA: 3011208 BP.

AB069245; AB069245;

29-AUG-2002 (first entry)

DE Listeria innocua DNA sequence #684.
 XX Antibacterial: Listeria; food contamination; mutational analysis;
 KW infection; ds.
 XX
 OS Listeria innocua.
 XX
 PN WO200228891-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-FR03061.
 XX
 PR 04-OCT-2000; 2000FR-0012697.
 XX
 PA (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Kunst F, Glaser P;
 XX
 DR WPI; 2002-332479/37.
 XX
 PT New genomic sequences from Listeria species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators -
 XX
 PS Claim 5; SEQ ID 2058; 180pp; French.
 XX
 CC The present invention relates to nucleic acid sequences
 CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
 CC and primers for identification and/or detection of Listeria (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in
 CC anti-Listeria vaccines.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SO Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;
 Query Match 53.2%; Score 205.8; DB 24; Length 3011208;
 Best Local Similarity 70.9%; Pred. No. 1.1e-46;
 Matches 273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
 Oy 3 TCCTCTCCCTCTCACCAGAAATGAGAAACACGTCATCTAACAGTCACCAAGACCTAAA 62
 Db 2066305 TCCTCTCTCTCTTTTGGAAAATAGCAGACGATCAATCACCACGCCCTAGT 2066364
 Oy 63 CCATGTGACCCGATATCGGATACGGTTCCACAAATCCCTAATAGAAATCATACACG 122
 Db 2066365 CCATGAGAACCAAAATTTGATACGGCTCCACAAACCAAGAGATAAAGTCATAATC 2066424
 Oy 123 TCTGTACGCAATTTCCATATTATCTACTTGTTAACGCTAATACGACGGTTTTAGAT 182
 Db 2066425 TGATCAACGCAATTTCTGGTATTCTACTTATTAATCGTAACAAATGTTTATATGAC 2066484
 Oy 183 TTGTATATAATTTGACGACCATTTTCATCGCTTTGTGTCAATCTTCACGACCTTAAC 242
 Db 2066485 CGGTAAAGAAATTTTGTACTTGTTCGTCGATCGGTAACCCCTTCACACCATTTGTA 2066544
 Oy 243 ATTAATAATATAACATCCGCTTCATCATGCGATTTCTCGCTGCGCTAATTTGTGTT 302
 Db 2066545 ATTAATAATATAATACGCTCTGCTTATCATTCGATTTCCGCTGTGCGCAATTTGCTCT 2066604
 Oy 303 TCGAATGTGTCATACCAATTTCAATACACGCTGATCAATTAATTTGAATCATGTGTT 362
 Db 2066605 AAGAAATGCTGCTCGGAAATCAATACACGCTGATCAATTAATTTGTTTCCA 2066664
 Oy 363 AACCATTCCTGAAGATAATATAC 387
 TTTTTT TTTTTT TTTTTT TTTTTT

Db 2066665 AGCCATTCGCGTGAATTATATATGCG 2066689
 RESULT 15
 ABK75008/C
 ID ABK75008 standard; DNA; 1311 BP.
 XX
 XX
 AC ABK75008;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #2299.
 XX
 XX Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 OS Bacillus licheniformis.
 XX
 PN WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 05-OCT-2001; 2001WO-US31437.
 XX
 PR 06-OCT-2000; 2000US-0680598.
 XX 27-MAR-2001; 2001US-279526P.
 XX
 PA (NOVO) NOVOZYMES BIORECH INC.
 XX (NOVO) NOVOZYMES AS.
 PA Berka R, Clausen IG;
 XX
 DR WPI; 2002-416684/44.
 XX
 PT Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 XX
 PS Claim 4; SEQ ID NO 2299; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequence tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 1311 BP; 376 A; 279 C; 353 G; 303 T; 0 other;
 Query Match 45.1%; Score 174.4; DB 24; Length 1311;
 Best Local Similarity 67.1%; Pred. No. 6.1e-39;
 Matches 24; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
 Oy 20 AAAATGAGAAACACTGCACTTAACAAGTCACCAAGACCTAACAACATGTGACCTGATAT 79

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:09 ; Search time 1580.02 seconds
(without alignments)
10020.134 Million cell updates/sec

Title: US-09-815-242-1463

Perfect score: 387

Sequence: 1 gatctctctctcttcacca.....ttcactgaagaataaac 387

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl:.*
1: gb_da.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_hum.*
32: em_htg_inv.*
33: em_htg_mus.*
34: em_htg_pla.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vfl.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	387	100.0	1308	6	AX622668
2	387	100.0	301550	1	AP003134
3	387	100.0	333750	1	AP004827
4	387	100.0	346900	1	AP003362
5	262.2	67.8	3269	1	AE270032
6	262.2	67.8	3269	6	AX145350
7	262.2	67.8	300892	1	AE016747
8	210.6	54.4	290117	1	AE017028
9	209	54.0	304680	1	AE017002
10	205.8	53.2	313450	1	AL596170
11	205.8	53.2	319630	6	AX413016
12	205.8	53.2	347050	1	AL591981
13	205.8	53.2	349980	6	AX417046
14	205.8	53.2	349980	6	AX641670
15	198.4	51.3	24887	1	BACSERA
16	198.4	51.3	213680	1	BSUB0012
17	191.4	49.5	302173	1	AE016951
18	188.2	48.6	300050	1	AP004599
19	185	47.1	302050	1	AP001512
20	182.4	47.1	302050	1	AL935257
21	174.4	45.1	1311	6	AX433884
22	162.6	42.0	11071	1	AE006309
23	162.6	42.0	12434	1	AE006498
24	162	41.9	1311	6	AX607165
25	162	41.9	44145	6	AX602195
26	162	41.9	174050	1	SAG766852
27	160.4	41.9	20601	1	AE014265
28	159.4	41.2	52276	1	AE014141
29	159.4	41.2	323825	1	AP005146
30	157.8	40.8	12370	1	AE009978
31	153	39.5	12540	1	AE008523
32	151.4	39.1	1308	6	AX570314
33	151.4	39.1	5066	6	AX218847
34	151.4	39.1	5066	6	BD003759
35	151.4	39.1	10310	1	AE007464
36	151.4	39.1	151947	2	SPNEU1902
37	151.4	39.1	349980	6	AX571764
38	151.4	39.1	349980	6	AX571765
39	149.8	38.7	13860	1	AE015016
40	148.2	38.3	3737	1	AB016077
41	139.4	36.0	301278	1	AE015939
42	125.2	32.4	296750	1	AP003191
43	117	30.2	10861	1	AE007680
44	109.2	28.2	3557	1	AY094626
45	101	26.1	960	6	AX144037

ALIGNMENTS

RESULT 1
AX622668/c
LOCUS AX622668 5631 from Patent WO02094868. 1308 bp DNA
DEFINITION AX622668
ACCESSION AX622668
VERSION AX622668.1 GI:28450653
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Masignani,V.C., Mora,M.C. and Scarselli,M.C.
Staphylococcus aureus proteins and nucleic acids
Patent: WO 02094868-A 5631 28-NOV-2002;
Chiron Spa (IT)

FEATURES
source Location/Qualifiers
1..1308
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/db_xref="taxon:1280"

BASE COUNT 450 a 184 c 278 g 396 t

ORIGIN

Query Match 100.0%; Score 387; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCTCTTCCTCCCTTCACCAAAATGAGAAACACATGTCATCTACAGTCACCAAGACCTA 60
|||||
Db 508 GATCTCTTCCTCCCTTCACCAAAATGAGAAACACATGTCATCTACAGTCACCAAGACCTA 449
|||||

QY 61 AACCATGTGACCCGTATATCGGATACGGTTACCAAAATCCATATGATAGAAATCATACA 120
|||||
Db 448 AACCATGTGACCCGTATATCGGATACGGTTACCAAAATCCATATGATAGAAATCATACA 389
|||||

QY 121 CGTGTGACGATTTTCATATTTACTTTGTAAACGCTAATACGACCGGTTTTTTAG 180
|||||
Db 388 CGTGTGACGATTTTCATATTTACTTTGTAAACGCTAATACGACCGGTTTTTTAG 329
|||||

QY 181 ATTTGTATTAATAATTGACGACCATTTTCATCGCTTTGTGTCAATCCTTCACGACGTTAA 240
|||||
Db 328 ATTTGTATTAATAATTGACGACCATTTTCATCGCTTTGTGTCAATCCTTCACGACGTTAA 269
|||||

QY 241 CCATTAATAATAATAATACATCCGCTTCATGCGATTTCTGCTGCGCTTAATTTTGG 300
|||||
Db 268 CCATTAATAATAATAATACATCCGCTTCATGCGATTTCTGCTGCGCTTAATTTTGG 209
|||||

QY 301 TTTGGAATGTGTCATCACCATTTCATATACACCGTATCATATTAATTGAATCATGTG 360
|||||
Db 208 TTTGGAATGTGTCATCACCATTTCATATACACCGTATCATATTAATTGAATCATGTG 149
|||||

QY 361 TTAAACATTTCACCTGAAGAATAATAC 387
|||||
Db 148 TTAAACATTTCACCTGAAGAATAATAC 122
|||||

RESULT 2
AP003134 301550 bp DNA linear BCT 24-APR-2003
LOCUS Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
DEFINITION genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE Bacteria; Firmicutes; Bacillales; Staphylococcus.
AUTHORS
1
Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Iian, J., Ito, T., Kanamori, M.,
Matsunaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,
Mizutani, O., Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C.,
Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J.,
Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,
Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.
TITLE Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 301550)
AUTHORS Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,
Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2001) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center, 2chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,

COMMENT
FEATURES
source
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
On Jun 12, 2001 this sequence version replaced gi:13701258.
Location/Qualifiers
1..301550
/organism="Staphylococcus aureus subsp. aureus N315"
/mol_type="genomic DNA"
/strain="N315"
/sub_species="aureus N315"
/db_xref="taxon:15879"
complement(249..947)
/gene="SA1293"
complement(249..947)
/gene="SA1293"
/note="ORFID:SA1293"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42553.1"
/db_xref="GI:13701259"

gene
CDS
complement(1001..1588)
/gene="SA1294"
complement(1001..1588)
/note="ORFID:SA1294"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB4254.1"
/db_xref="GI:13701260"

gene
CDS
complement(1578..2153)
/gene="SA1295"
complement(1578..2153)
/note="ORFID:SA1295"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB4255.1"
/db_xref="GI:13701261"

gene
CDS
complement(2167..3411)
/gene="SA1296"
complement(2167..3411)
/note="ORFID:SA1296"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB4256.1"
/db_xref="GI:13701262"

gene
CDS
complement(3418..4716)
/gene="arocA"
complement(3418..4716)

gene
CDS
complement(3418..4716)

gene /gene="aroA"
/note="ORFID:SA1297"
/codon_start=1
/transl_table=11
/product="3-phosphohikimate 1-carboxyvinyltransferase"
/protein_id="BAB42557.1"
/db_xref="GI:13701263"
/translation="MVSEQLIDISGPLKEIEVPDGSMTTHRAIMLASLAEGSTNIX
PLIGEDCRRTMDIFRLGVDIKEDOKLVNNSPGYKAFKPHQVLYVNGSGTTRLLA
GILSGIESVSGDYSIGKRPMDVRLRLKMDANEGIEDNNTPLITKPSYIKGN
XOMEVASAOKSAIILFASLSFNNTTYIKELDVSRNHTETFRHNPIDEMERISITTT
PDAIHIKPADFHPGDISAFAFIVAALTPPSDVITINVGINPRSGIIDIVERKG
GNIQLFNQTTGAPPTASIRIQYTPMLQPTIEBELPKAIDELPVIALICTQAVGST
IKDAEELKVEETNRIDTADMLNLGFELOPTNDGLIHPSEFKTNAIVDSLDRIG
MMLAVASLSESPVKIKOPDAVNVSPFGFLPKLLENEG"
complement(4726..5790)
gene /gene="arob"
complement(4726..5790)
CDS /gene="arob"
/note="ORFID:SA1298"
/codon_start=1
/transl_table=11
/product="3-dehydroquinate synthase"
/protein_id="BAB42558.1"
/db_xref="GI:13701264"
/translation="MKIQTPSPNNYPIFEVHGAIDHISTYIDQFDSFILDEHVNQ
YFADKFDLISYENHVKVILPAGEKTFEYOETLEYILSHVTRMTATIAVGGAT
GDFAGVYATLLRGVHPHIOVPTIILHSDSVGKVGISKOGNLICAFRRPAVYID
LDFKTLRPEQILISGVAEYKHALNGESTTOEIQHFKRBEILQISNGDKTIKAI
ERKLDIVADEKQGVKFLNLGHTFGHAEVYHKKIAHGAVMIGIILYQIVANILEN
SNHDIGHYINLYKRGYPLETITIDIDETIYOYMLSDKNDKOGVOMVLIKHFQDIYV
QHIDQTLQHACQOLKTYEK"
complement(5816..6982)
gene /gene="aroc"
complement(5816..6982)
CDS /gene="aroc"
/note="ORFID:SA1299"
/codon_start=1
/transl_table=11
/product="chorismate synthase"
/protein_id="BAB42559.1"
/db_xref="GI:13701265"
/translation="MRLTSGESHGPOLVIVGIPANLEIKVEDINKENFKRGGYG
KGRHMOIEXDTEIVISGVNRYGSPITFVWVNDPTHRKINGADPISEEEENK
RTTKPRPGHADDVFGMKYNHRDLRLYESSARETAAVAVGLCVLLOODIDY
SRVYEIGIKDKDFYDSETEKALMDRDVVIDSLAQARDKIDEKKNESGDIGYV
QYVENNRPVGVGSYHTRDKLDKINAGVINSIAFKVSGEGFKAERKSGSEIODEI
LVSEIGYRGSNHLGLEGSGNSGMPTIVNGVAKPIPTLYKPLNSVDINTKEDFKAT
IERSDSCAVPAASIVCEHVVAFEIARALDEFOSNHIEOLOQOITADRRQLNVEF"
complement(7368..7568)
gene /gene="SA1300"
complement(7368..7568)
CDS /gene="SA1300"
/note="ORFID:SA1300"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42560.1"
/db_xref="GI:13701266"
/translation="MKNITISILGIIMFLKLMELLYGAIFLDRPLNDITKIFLTLI
YIPYLVAKELIIFLKSYNKSA"
complement(7597..7722)
gene /gene="SAS045"
complement(7597..7722)
CDS /gene="SAS045"
/note="ORFID:SA5045"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42561.1"
/db_xref="GI:13701267"
/translation="MWLVMIINHKARFLCCSKLLEKYSFLVTRIKKRGIFIK"
complement(7777..8226)
gene

CDS /gene="ndk"
complement(7777..8226)
/gene="ndk"
/note="ORFID:SA1301"
/codon_start=1
/transl_table=11
/product="nucleoside diphosphate kinase"
/protein_id="BAB42562.1"
/db_xref="GI:13701268"
/translation="MERTFLMKRPDANVRNLIGEVISRIERKGLKLVGKLMQVPMEL
AETHYGHGQKRPYNDLISFTSAFVFAVVEGDVAVSRHIIIGSTNPSEASPSIR
GDGLITVGRNIHGSDSLSEASEREINLMFNENETTSYASPRDMLYE"
complement(8318..9277)
gene /gene="gerCC"
complement(8318..9277)
CDS /gene="gerCC"
/note="ORFID:SA1302"
/codon_start=1
/transl_table=11
/product="heptaprenyl diphosphate synthase component II"
/protein_id="BAB42563.1"
/db_xref="GI:13701269"
/translation="MKLNMNNEIKVEQRLEKAIKSKDSVLEQASLHLSSGGKRV
PAFVILSSQFGKDEQTSQRTYQVAVALLELHMATLVHDVIDKSDKRGRKLTTSKMD

Query Match 100.0%; Score 387; DB 1; Length 301550;
Best Local Similarity 100.0%; Pred. No. 1 2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCTTCTCTCTTCACCAAAATGAGAAACACGATCTACACGACCAAGACCTA 60
|||||
Db 13271 GATCTTCTCTCTTCACCAAAATGAGAAACACGATCTACACGACCAAGACCTA 13330
|||||
QY 61 AACCATGTGACCCCGATATGCGATGACGTTCCACCAATCCATAGATAGAAATCATACA 120
|||||
Db 13331 AACCATGTGACCCCGATATGCGATGACGTTCCACCAATCCATAGATAGAAATCATACA 13390
|||||
QY 121 CGCTGTACGATTTCCATATATCTACTTTGTTAAACCGCTAATACGACGGTTTGTAG 180
|||||
Db 13391 CGCTGTACGATTTCCATATATCTACTTTGTTAAACCGCTAATACGACGGTTTGTAG 13450
|||||
QY 181 ATTTGTATAAATTTGAGCGACATTCATCGCTTGTGTCATCTTCACGACGTTAA 240
|||||
Db 13451 ATTTGTATAAATTTGAGCGACATTCATCGCTTGTGTCATCTTCACGACGTTAA 13510
|||||
QY 241 CCATAAAAATATATACATCCGCTCATCTATGCGGATTTTCGCTGCGCTTATTTGTG 300
|||||
Db 13511 CCATAAAAATATATACATCCGCTCATCTATGCGGATTTTCGCTGCGCTTATTTGTG 13570
|||||
QY 301 TTTCGAATGGTGCATTCACCAATTCATACACCGCTGTATCAATATTTGAATCATGTG 360
|||||
Db 13571 TTTCGAATGGTGCATTCACCAATTCATACACCGCTGTATCAATATTTGAATCATGTG 13630
|||||
QY 361 TTACCATTCACCTGAAGAATAATATAC 387
|||||
Db 13631 TTACCATTCACCTGAAGAATAATATAC 13657
|||||

RESULT 3
AP004827 333750 bp DNA 1linear BCT 24-Apr-2003
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 6/10.
ACCESSION
AP004827 BA000033
VERSION
AP004827.1 GI:21204509
SOURCE
Staphylococcus aureus subsp. aureus MW2
ORGANISM
Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1 Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
Negai,Y., Iwama,N., Asano,K., Naito,T., Kuroda,H., Cui,L.,
Yamamoto,K. and Hiramatsu,K.

TITLE	Genome and virulence determinants of high virulence
JOURNAL	Community-acquired MRSA
PUBLISHED	Lancet 359 (9320), 1819-1827 (2002)
REFERENCE	22040717
AUTHORS	2 (bases 1 to 333750) Director-general, Biotechnology Center, Aoki, K., Oguchi, A., Nagai, T., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE	Direct Submission
JOURNAL	Submitted (06-MAR-2002) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2chome 48-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bioente.go.jp, URL:http://www.bio.nite.go.jp/, Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES	Location/Qualifiers
source	1. 333750 /organism="Staphylococcus aureus subsp. aureus MW2" /mol_type="genomic DNA" /strain="MW2" /db_xref="taxon:196620" complement(278. . 619) /gene="MW1341" complement(278. . 619) /gene="MW1341" /note="ORFID: MW1341" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB95206.1" /db_xref="GI:21204510" /translation="MRTNTEKEVFNLEQEDIVFAKANKTLGSEAKPYLDEYHSKIIDD FKVNDITGINDFDLNOYPVPMFKERYDMIERKHYFMGYRQMKTFKELLKMA SYQRLKNKOY" complement(624. . 1283) /gene="nth" complement(624. . 1283) /note="nth" /gene="nth" /note="ORFID: MW1342" /codon_start=1 /transl_table=11 /product="endonuclease-like protein" /protein_id="BAB95207.1" /db_xref="GI:21204511" /translation="MWSKKALMIDVIANMPDAECLEKHDNPELTIAVLNAQCT DVLVAVRTTELFEKRYTPEEDYLAVSDELMNDIRSIGLYRNKAKNIKRLCOSLIDQFN GEIDPTNKELESLAGVGRKTANVMSVAFDEPSLAVDTHVERSVSGINRMKDVNQ VEDRLCSVIPBDRMNRSHOQLFEGRYCLARKPKCDICPLEDCREGQKRYKASLKE A" complement(1273. . 1959) /gene="MW1343" complement(1273. . 1959) /gene="MW1343" /note="ORFID: MW1343 hypothetical protein, similar to chromosome replication initiation protein dnad" /codon_start=1 /transl_table=11 /protein_id="BAB95208.1" /db_xref="GI:21204512" /translation="MDKYQLKAPVYIRRELLDHYSDGLDEODLVILLKLIYASETS NKPSIELQKSGMQRDITWVIONLLORELELQVQKDEGRTEYFMNDPFEKL SHLQOSMETEONSKERKQLEFVLDYFARPLSPYEITLQWIDVDHDSATIQ AALDEANSLNLSFKYMDRILLNMKKNVKTIDSRKIREKFNKPKMHTVTKPKFD WLNGENLDGK" complement(2287. . 3579) /gene="asnS" complement(2287. . 3579) /gene="asnS" /note="ORFID: MW1344" /codon_start=1 /transl_table=11 /product="asparaginyl-tRNA synthetase"
gene	/protein_id="BAB95209.1" /db_xref="GI:21204513" /translation="MKTIIKAKDHLNODVTIGAWLTNRSSGKIAPLQLRDGTGFMQ GVVAVSEVDEEVEFLAKEITQESSLYVTGITEEDNRSDLGEMQKSLIEVSEADYP ITPKNGHTEFLMDHRLMLRSKOHAVMKIREVYRATYEFNKFPKVPDPLTAS APEGTSLEPHRYKFPDPAFISQSGOLYLEAAMAMGAKVPSGSPFRASKSRRLIE FMWIGEMAFTHAESAELQEDYVTHVYKSVLENKLEKITERPSTLKEKYANFPFR ISYDAIEFLASSEGFDLEMGDEGAPHELTAAINYDLVPLFTNTPTKIKPKYQPNP ENBEETVLCADLAPBEGEIIIGSESRVDELLEQVRKEHGLEDEEAYSYVDLRRYGS VPHCGFGGLGERTVAMISGVEHVRETAFFPRLNRLP"
gene	complement(3901. . 6594) /gene="ding" complement(3901. . 6594) /gene="ding" /note="ORFID: MW1345" /codon_start=1 /transl_table=11 /product="probable AMP-dependent DNA helicase ding" /protein_id="BAB95210.1" /db_xref="GI:21204514" /translation="MGMAVYAVDLETTGNOLFDDDIIOIGITPVFNNOIIDTYHSMI RNLIEIPFIOALTSIEENMLQAPYFNQAOEIDKIKDCIFVHANYDFDNLTKKA FKDCNIQYRPPKVIDTLEIFKIAFPYDSYSELAEAHGITALAHARADEDAATTA LMILAEKPEKPLDITLKOIYLSKQLYDLDYDIFEEVROYDAKPLKSEKPEOIT YRKQYDFKPTTNVNGSLKSLAKVADOLGITYRQOYLAEITIDQLMHSKAMIEA SLGSKSLAYLALALMNIETGKHVISTNRKLDQSULEKDIYPMNARLAFKIAL IKSKSDYISLGLISQILKDDTSNVEVNLKQOLLITETPEBGDIQELNLGQGMFY DQKIETVYPARHDVYVNFIRKNAONIGITNHNKLAHSDEVNSIYOLFDCIYDEA HRLPYALNOVYNELSYADIKYQGLIGKNEKEKLLKALIDOLEKRIIEKDIAPIDI FGKASMEIHLEINOLFSTETITINDSGVDDDIHRPHNYFTEETKQILNDLHAIID KLNTLEIFNGISHTKYSLSKROLLYXLDKRNKEIOSLKAGHTSPISIKNLSSKSTR LTVQYAVKDVLTQVLEKFSLEITSGTLRFNHSFEAKQLEKNDVHRTVEVNTSL QSAKNTSVFIPSDVASYOYKNIDEVASIVSIIETTYTSKCLVETSYKMMHMQ DMLNELPEFEDYVVLTOQONONYKIVQOFNNEDKAILLGTSTFFGEPFOANGIKCMV IAKLPFKMHNAAKYWLMDSEFTSEFTSEYVLPDAVYRFRQGLIRSENDGIIYSFD DRLINSYKKNFEQFTLENYROKKDIOQFGKLLRQIQKKKK" complement(6618. . 7589) /gene="MW1346" complement(6618. . 7589) /gene="MW1346" /note="ORFID: MW1346" /codon_start=1 /transl_table=11 /product="truncated hypothetical protein" /protein_id="BAB95211.1" /db_xref="GI:21204515" /translation="MSKYSQDVLQLLYKKNPNYISGOSIABSLNISRYAVKVIDOLK LEGCRIDSVNKHGHLLOQLPDLYOGIIIDQYTKSSALDFSEVYDSIDSTOLAASKL VGNOSSEFLLSDEOTKGRGFRMRHSSSGGGLMWSVYLRPVASMSIKNPLFALG IRDAIQHFSQDEPVKYPNDIYIDNGKSGCLTEWYVANNDDIGALIGGINTLQOLE NFDESIRHRATSVQLHDKNKLDROYFLERLQLEIKRKNQFLTLPFSEIREBYIAASN IWNRTLLFTENDKQFGQALDLDYDGYLIVDEAGESHRLISADIDF" complement(7576. . 8778) /gene="MW1347" complement(7576. . 8778) /gene="MW1347" /note="ORFID: MW1347 hypothetical protein, similar to poly(A) polymerase" /codon_start=1 /transl_table=11 /protein_id="BAB95212.1" /db_xref="GI:21204516" /translation="MDKSLFEOARPILEQIODNGFEAVYVGSVRDYVMGRNHDIDI TTSAPDEIESIFSTPIVGRKEHGTINVPDENDEVTTPFAEEDYVDRHPSGVTV RDLYEDLORDRDTMNAIAMRAYKLYDFEQODINNRILRTVGAIEBEPEDALAMI RCLAROSOLSPDIATPEARMARIOMADIKLISERIVLETKRMGINVESKFNLS LKAEVYMPFELDMNOINVTAEIDELLIAIVSKPDIYSLPKLISNOVDIND YIOINNALPSITTKEQLKMPYVDYDTHLIKRVWVAADYIKANDIOGHEPLVYNQOTID ETLHRLPMPHNRDMWNGVLMHLNANSGWLDVLEQLQIATVYTKVSNSETEILK WVDHVKI"
gene	complement(8783. . 9925) /gene="MW1348" complement(8783. . 9925)
CDS	

```
/gene="WM1348"
/note="ORFID:WM1348
hypothetical protein, similar to lipopolysaccharide
biosynthesis-related pr homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517"
/translation="MKIGTCYPSMGSGIATETGIKLAERGHVHTTSNIPRIR
KLPIMIFHOVEONVAYFOYPPYDITSTKTAERYKEGIDLIMHYAPPAICIL
AREMSGRDKIKITTIHGTDITVLGDSHLOQAIKGIKESDVIYTSVSLADETHIIT
ETNKEIPIYVNEPENEPTKHTALKSQFGLAPDEKVLIHVSNRYQKRIIDTIIETE
AKVREKIPSKLLDGDPELPMRQTLKELNEVDVLELGKQDCSEFYOLSDIVLL
SEKSEFGLTEAMKTVGPVIGSNAGIKREVIKHEGTGVVDGDCSDSYAIRLLE
DKVLYNKLQKNMLADIARFSESLITDQVEYYVQKMLNEHNSKGE"
complement(10167..10484)
/gene="WM1349"
complement(10167..10484)
/gene="WM1349"
/note="ORFID:WM1349"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translation="MKSVMEMOREVDEYIGQPKTGFSPPLANLARTEVEGLAREIN
HTYGEKKKKDSEANTIKAEIGDNLFVLLCLANSNGIDMTESFNETMEKFNTRDKNRF
ERK"
complement(10820..11518)
/gene="WM1350"
complement(10820..11518)
/gene="WM1350"
/note="ORFID:WM1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translation="MSSIFMTIYFVILMLPLMAOHKYSNVEKYSOVSRSCTGR
EVALEILANGIYDEYVKGSGFLDHDPRKKYVLSLPANDRPSVAGTALAHBEVG
HAIDHOGYWFLEPRFALVPAVNLGSSLYSTMILMIGILITLALGSAFGSALMIGAGLM
```

```
Query Match 100.0%; Score 387; DB 1; Length 333750;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

gene
CDS
1 GATCTTCTCTCTCTACCAAAATGAGAAACAACTGCATCTAACAAAGTCACCAAGACCTA 60
|||||
23844 GATCTTCTCTCTCTCTACCAAAATGAGAAACAACTGCATCTAACAAAGTCACCAAGACCTA 23903
|||||
61 AACCATGTGACCCGTATATCGATACGGTTCACCAAAATCGTAAATGGAATCATATA 120
|||||
23904 AACCATGTGACCCGTATATCGATACGGTTCACCAAAATCGTAAATGGAATCATATA 23963
|||||
121 CGTCTGTACGCAATTCATATATCTACTTTGTTAAACGCTATACGACCGCTTTTGTAG 180
|||||
23964 CGTCTGTACGCAATTCATATATCTACTTTGTTAAACGCTATACGACCGCTTTTGTAG 24023
|||||
181 ATTTGTATTAATTTGAGCGACCAATTCATCGCTTTGTGTCAATCCTTCAGCAGCTTAA 240
|||||
24024 ATTTGTATTAATTTGAGCGACCAATTCATCGCTTTGTGTCAATCCTTCAGCAGCTTAA 24083
|||||
241 CCATATAAATATATACATCCGCTTCATATGCGCATTTTCGCGCTTAATTTGTG 300
|||||
24084 CCATATAAATATATACATCCGCTTCATATGCGCATTTTCGCGCTTAATTTGTG 24143
|||||
301 TTTGGAATGTGCATACCAATTTCAATACCACTGTATCATTAATATGGAATCATGTG 360
|||||
24144 TTTGGAATGTGCATACCAATTTCAATACCACTGTATCATTAATATGGAATCATGTG 24203
|||||
361 TTAACCATTTACCTGAGAAATATAC 387
|||||
24204 TTAACCATTTACCTGAGAAATATAC 24230
|||||
```

```
RESULT 4
AP003362
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
sequence, section 5/9.
AP003362 BA000017
AP003362.2 GI:14247083
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
Staphylococcus aureus subsp. aureus Mu50
Staphylococcus aureus subsp. aureus Mu50
Bacteria; Firmicutes; Bacillales; Staphylococcus.
1
Kurida,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)
21311952
11418146
2 (bases 1 to 346900)
Ohta,T.
Direct Submision
Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
of Medical Technology and Nursing, Department of Medical
Technology, 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan
(E-mail:tohta@tsukuba.ac.jp, Tel:81-298-53-3454,
Fax:81-298-53-3454)
On May 29, 2001 this sequence version replaced gi:13875626.
Location/Qualifiers
1..346900
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
12..263
/gene="SAV1313"
12..263
/gene="SAV1313"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57475.1"
/db_xref="GI:14247084"
/translation="MULTIKLKEDEQVYIYETIPEDISNGKGSYTFNKKDAEVIDFS
LSEIENEYFMYLRNKSFSVVRDFLEKQEPENYKIATW"
987..1181
/gene="SAV1314"
987..1181
/gene="SAV1314"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57476.1"
/db_xref="GI:14247085"
/translation="MMLQSYIGIKLVKLTNNKILIGKVIYIDPDKVDNPDGNSIEID
TRITVDISENKIKITVLRK"
complement(1725..2750)
/gene="SAV1315"
complement(1725..2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"
```

gene
CDS
3020..3217
/gene="SAV1316"
/gene="SAV1316"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB5748.1"
/db_xref="GI:14247087"
/translation="MEQIKLFTFTAEELLEKNINAFLSSEATNMLKVNITKEIE
ERFPNNEEFNALILTSYK"
3238..4755
/gene="SAV1317"
3238..4755
/gene="SAV1317"
/codon_start=1
/transl_table=11
/product="cardiolipin synthetase homolog"
/protein_id="BAB57479.1"
/db_xref="GI:14247088"
/translation="MEFVPEFIEGKTMRYTFSNDLGTFTIILAIGFIINLVAFI
FLERNRTASTAMLEVLVLELIGFTIYLPFGRTVSARKLNNGANVLDFDGLK
OOIESPDKNGTGNKOKOHHDIVSMLEDDOGLTENNKHDPHTDGLNDLYQVLD
IKNAKEYIHLEYTPRALDGLKRIHLEBKILIDVDGSKNKKMNFPE
KSLGGEVNEFASKLPFLNFRNNRNHRIIVIDGOLGVGGNIDEXVLGKRLIYW
RDTHLRIQDAVDALQRLFLDMNSQAHHPQFEYDVKYPKNGPLGNSPIQAAAGP
ASWMOIEGYGTAKMISAKSVYLOSPLYIPDINSYINAIKAKSGVDHAIAMP
HPLVYATFENASDLSSGVKIYTYENGFTLSKMCILIDDEIYSGVTANDERSFELN
EVNAFYVDENLAKDLVAAYEHDITKSKQLTKESYANRPLSVKFKESLAKVSPYL"
4939..5436
/gene="SAV1318"
4939..5436
/gene="SAV1318"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57480.1"
/db_xref="GI:14247089"
/translation="MIOISNINSEFNKRCVLKNISPDIEGKCIALIGKNAGKSTLI
DLIGVNAHNSGEIFDKDLLOSERNISIMOKTMRPDOLKIEIINLYXSPYENPL
EETIELTKDSSQLNPFVNLKSGGQRLDPVLSLIGPQLILDEPTSMOIEIREY
PMSIT"
5461..5838
/gene="SAV1319"
5461..5838
/gene="SAV1319"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57481.1"
/db_xref="GI:14247090"
/translation="MITYSHYIEFVRMSDKIILINGEIIILNDSTSHIRTNQOSIT
LSDYIRKLKLDLVLQKNHNGTIKITSNVNPTILYLOOLHMLDIEIOKXSVI
DSYFNKKORGSGSYDFKLEENRI"
5807..6538
/gene="SAV1320"
5807..6538
/gene="SAV1320"
/note="similar to ABC transporter integral membrane
protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57482.1"

gene
CDS
6542..7633
/gene="SAV1321"
6542..7633
/gene="SAV1321"
/note="similar to two-component sensor histidine kinase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57483.1"
/db_xref="GI:14247092"
/translation="MKELKDTSAEISLTYLIFPFIAGIFPNVGPKKLYIIVTF
SLSYLLVLYNNRLNLTMLTYILIHIFPIICFVESVHPLSLFFSYSAFAPFTK
NVKRTANLEFILTMICTIITTYLLYNNYFVAMVYVYLSLMDNFKKMKREYKE
IAKNRHINTLAEORHRIGODLHDTGVFASLSKSELAKLIDTVEKAKBL
AINKLSRESLNKREIIDDVKLPSPTEIDSIKRVLKADIDTFENKELAOVSPK
OSMLVITREATINNVIRKANASKVHGKLTAVNNHKLMLMIEDDGKIDSDCEKYSISO
RVQHLNGTIAVDSTNGTKIIEISTGIA"
7630..8232
/gene="SAV1322"
7630..8232
/gene="SAV1322"
/note="similar to two-component response regulator"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="MFSLIAEQNNLRQAMVOLIKHGDFEILADTNGDLAMKLI
EYNPVNIYDIEMPGMTGLEVLAEIRKHLNIVITVTFKRPGEYKAVNDVDAV
IKRSIEBELYERINKVNGEKEYSATLMSFPYDKRPLPKQIYVIREIGNGLSKEI
SEKFLITDGVTRNTYIYIDKLPADNRFPDAKANKGMI"
8672..9205
/gene="SAV1323"
8672..9205
/gene="SAV1323"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNISNENVTYLLAVIIAIIIGFIQIFKPLIYTAVISILGII
FVGEITYLVISFYNNKRN"
8672..9205
/gene="SAV1324"
8672..9205
/gene="SAV1324"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"

Query Match 100.0%; Score 387; DB 1; Length 346900;
Best Local Similarity 100.0%; Pred. No. 1,2e-77;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCTTCCTCCCTTCACCAAAATGAGAAACACGTCAGCTACAGTCACCAAGACCTA 60
|||||
Db 200303 GATCTTCCTCCCTTCACCAAAATGAGAAACACGTCAGCTACAGTCACCAAGACCTA 200362
|||||
Qy 61 AACCATGTGACCCGTGATATGCGATTACGCTTACCAAAATCCTAATGATAGAAATCATPACA 120
|||||
Db 200363 AACCATGTGACCCGTGATATGCGATTACGCTTACCAAAATCCTAATGATAGAAATCATPACA 200422
|||||
Qy 121 CGTCTGACCATTTCCATATTAATCTTCTGTTAAACCGCTAATACGACCGGTTTTTAA 180
|||||

121 CGTCTGACGCAATTCATATATCTACTTGTGTAACCGCTAATACGACCGGTTTTTAG 180

361 TTAACCATTCACCTGAAGATAAATAC 387
 ||||| | | | | | |||||

Db	2657	TTACCACTCGCCAGATGAATAATAC	2683
RESULT 7			
AE016747			
LOCUS	300892 bp	DNA	linear BCF 01-Jan-2003
DEFINITION	Staphylococcus epidermidis ATCC 12228, section 4 of 9 of the complete genome.		
ACCESSION	AE016747	AE015929	
VERSION	AE016747.1	GI:27315369	
KEYWORDS			
SOURCE	Staphylococcus epidermidis ATCC 12228		
ORGANISM	Staphylococcus epidermidis ATCC 12228		
REFERENCE	Bacteria: Firmicutes; Bacillales; Staphylococcus. 1 (bases 1 to 300892)		
AUTHORS	Zhang, Y., Ren, S., Li, H., Yu, G., Lu, L., Lu, G., Jia, J., Tu, Y., Qin, Z., Chen, Z. and Wen, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-NOV-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China		
FEATURES	Location/Qualifiers		
source	1..300892		
	/organism="Staphylococcus epidermidis ATCC 12228"		
	/mol_type="genomic DNA"		
	/strain="ATCC 12228"		
	/db_xref="ATCC:12228"		
	/db_xref="taxon:176280"		
gene	113..346		
CDS	/locus_tag="SE0907"		
	113..346		
	/locus_tag="SE0907"		
	/codon_start=1		
	/transl_table=11		
	/product="acyl carrier protein"		
	/protein_id="AA004504.1"		
	/db_xref="GI:27315370"		
	/translation="MENPKVVDIYDRGVDADKYTEDAFKDDIGASDLDAELVM ELEDGFTIPDEAEKINTVDAAVYINLEK"		
gene	461..1198		
CDS	/locus_tag="SE0908"		
	461..1198		
	/locus_tag="SE0908"		
	/codon_start=1		
	/transl_table=11		
	/product="Rnase III"		
	/protein_id="AA004505.1"		
	/db_xref="GI:27315371"		
	/translation="MANOKKKEWVNHFOOKFTDDKMSLGIHFKNIDLVOOAFSSHSFI NDEPNRLFHNRLERLEIDAVLELTVSRFLPDRPHLPFGNLTAKMATIYCEPSTIYF ANKIKINELILGKEBEKTRPSLISDAFEAFYALYLDGDLDSVTFPAKVIAP YVEDELVGVDVFTQFOEYVHSQNGDVYOLIKKEGPAHRLFTSEVILEKAAVE GKGTKESEQAAEQAVKLMKNKSL"		
gene	1325..4894		
CDS	/locus_tag="SE0909"		
	1325..4894		
	/locus_tag="SE0909"		
	/codon_start=1		
	/transl_table=11		
	/product="chromosome segregation SMC protein"		
	/protein_id="AA004506.1"		
	/db_xref="GI:27315372"		
	/translation="WYLSIDAIGKSPADHTNVQDKVTAIYVNGSGKSNTDA IKWVLEQSAKSLRGAKMEDIIIFSGAHRKAQNFQAVKILDNHOSKLOIDSELVYT RLISGSESEYVNLNDRARLKDIIDFLDSLGKEAFSIISGRVDEILNARKINRQ IIEESAGVLYKKRKRAESIOLDHTEFNRPVEDIYDELEGRVEPLKEEAAIKRQ LSKMOSDVITVSDIDHYTEDNORLDRHLNHLKSOAEKREGOOAOIOLQYKXG RQONDYDIKLVNELYKATENYEQLSKLVNLEERKNOSERTNAREEEDNLNLESD STRNEQAQEKILAEILKNOKOLNKEVQDESLVITSDQHDKLEIKNSYTTLASE QSDVNDIDFLEHTINENAKSRSLDSRLEAFNQDKIDIONITQOKEQYSSKSGME KYEONIQOOLEQDLSKRLSEYENKLYOARENKLSRISLATOEDDYTYFFNGV KHLKAKDELRIGHGAVAEVINPSEMTOALETAGASLOHYVDNEKORAOIYOL KQRCGRATEFLPNTOPRHVADIKDVARSGGFTINASDAINVSAYKONITENLIG NTIYENLKHANELARAIYRIRIYITLEGDVYVPGGSMTGGAARKTKSILSODELST		

gene	MRNQLDYOQRTAFEFROFKEQTOAEQISEQYFSAQOYNLKEQVHHHELELRLK TDEAHLKNEHEEFERKNDGYQSDKSETKLKEQNHLIEIQDLKQLESDIENYITOLS KEGASTHQTOQLHQKOSDLAVYKERIKSOQVYRLDLKQSDSEROKIENVEKIKL FNSDEMGKDAFEKLEQIIOQOENVNQNLNQSLSEIKQORKDLNERKIEINESOLQKCH ODLSTENHYQODIKAKOSKLDVILNNAIDHNLNTYOLUTYERARERXDSPEDTNLARK VKITKMSIDELGFTNLNATBOPELNERYTEFNLEOFTDIREKETLEQIIEHMDKVE GRKTTFAVQDHFTTVFKOLREGGQAEILRTEDDYLKAGVDIYVPPGKKQQLHSLB SGERALSAIALLFALKVRSAPFVLDEVAALDEANVIRIAQYINLESTETQFTVI THRKGTEFSRDLRYGVTQMOESGVSKLVSNLNTIDVMEKEQA"		
CDS	4891..6117		
gene	/locus_tag="SE0910"		
	4891..6117		
	/locus_tag="SE0910"		
	/codon_start=1		
	/transl_table=11		
	/product="signal recognition particle"		
	/protein_id="AA004507.1"		
	/db_xref="GI:27315373"		
	/translation="MSFFKRLKDESSKSEDDIQKLDLETVDNSNANSDSNDPNSD EGYKPRKPKKLSADPDEGLISIDFEIETEAQIGAKFAGLERSKONFOQLNL IARYRVDEDFEALREMLITADVGNFVWKLDELRTEAORNIQETDRLREVYK IVEIYHOEDHSEAMNIEDGRLVNIVMGVNGVYKQTTIGKLAVRYQOEGKVMYLAG DFRAGAIQOOLNMGSRGVEYVSONEGSPAVVYDANAAKNKGVLDICTAGRL ONKSNIMQELDKMKRVINRAIPDAPEHALICLDATGONALSOARSEKFTVNSGVL TKLDGTAKGCTVLAIRNELHIPKYVGLGKMDLPFPNPESTVYGLPDMIQENDI PEETSRNSSVSEEGN"		
gene	6119..6451		
CDS	/locus_tag="SE0911"		
	6119..6451		
	/locus_tag="SE0911"		
	/codon_start=1		
	/transl_table=11		
	/product="conserved hypothetical protein"		
	/protein_id="AA004508.1"		
	/db_xref="GI:27315374"		
	/translation="MGONDLVKTIRMNLYLFEDFOSLITNKKQVLEFFYLQDSYSEI ADFFEYRQAVYINIRGTGDLVYDESKLRLLYOFERKRELYILMQSLNPBELLKQY ITQLEEL"		
gene	6482..7849		
CDS	/locus_tag="SE0912"		
	6482..7849		
	/locus_tag="SE0912"		
	/codon_start=1		
	/transl_table=11		
	/product="f1h protein"		
	/protein_id="AA004509.1"		
	/db_xref="GI:27315375"		
	/translation="NAFEGISDLQATFMOKMRGKGVTEADIKTMREVRLLAEADV NRYVVEPKVTVSERALGSDVMSQILPPGOVYIYIVQBELTSLMGENTSTKMYKNPPT VVMVVLQGGKTTAGKIALLRKRYNKPPLVAGDITRPARIIDOLQYTGKIDIP YVSGDVQSPQIVENALHAKKEHDLFVLIIDVAGDRIGALNNEQVEIKSPDEI MLVVDAMTGODAVNAQSPDEQDVGVTITKLDGDTRGALNISVYQKPIKFGM SEKLDGELTELFHPRMASRIIGMCDVSLTEKAAQODVDOEAKLDEKKMSSFTLDF LEOLDQVKNNGPLDDIMKPCMKKKGLDKLMSKQIDHIKAIIOISMTPKRSNPE TLVNSRKRRIAKSGRSLOEVNRLMKQFNDMKMMQFTGCGGKAKAKNQNMLKG MNLFP"		
gene	8281..8556		
CDS	/locus_tag="SE0913"		
	8281..8556		
	/locus_tag="SE0913"		
	/codon_start=1		
	/transl_table=11		
	/product="30S ribosomal protein S16"		
	/protein_id="AA004510.1"		
	/db_xref="GI:27315376"		
	/translation="MAVKIRLTPLGSKRNPFRIVVADARSPDGRILQLYGNPAH VNAPEVKIDBELAKLHDCAKFTDVYHNLSREGILKKFDEQKNK"		
gene	8683..9186		
CDS	/locus_tag="SE0914"		
	8683..9186		
	/locus_tag="SE0914"		
	/codon_start=1		
	/transl_table=11		

	Query Match	57.8%	Score 262.2	DB 1	Length 300892
	Best Local Similarity	79.8%	Pred. No. 2e-45		
	Matches 309	Conservative	0	Mismatches 78	Indels 0
				Gaps 0	
QY	1	GATCTCTTCCTCCACCAAAATAGAAACAAGTCATCTAACAAGTCACCAAGAAGCTA	60		
Db	296677	GATCTCTGANTCTTTATTTAAAGTTTCAACAACATGCAATCTGACAAATCTCCAAAGTCCTA	296738		
QY	61	AACCATGAGCACTGATATCGATACGGTTCACCAAAATCCTATGAATGAATCATACA	120		
Db	296737	ATCCATGTGAACACCAAAATAGAGATATGGATCTCCAAAGCCTTAAGATATGAATCATACA	296799		
QY	121	CGTCTGACGCAATTCATATTATCTACTTTGTAAACCGCTAATACGACCGGTTTTTAA	180		

Db	296797	TATCATTTAGCGATTTCACGATTATCAATTATTCACAGCTAAATACACAGGTTTCTTAG	2968
Oy	181	ATTTCGTATTAATTAATTTAGAGCGACCAATTTTCATCGCTTTGTGTCAATCTTCACGACGTTAA	240
Db	296857	ATTATATAAGACATTGTTCGCGACCAATTCGTCACCTTTGTGTAAAGTCCTTCCTACATTTGA	2969
Oy	241	CCATTAATAATAATATACATCCGCTTCATCTATGCGCATTTTCGCTCGGCTTAATTTGTG	300
Db	296917	CCATTAATAAATGATGATACATCTCTTCCTTCATTTGCTATTTTCGCTGCGACAGGATTTGAG	2969
Oy	301	TTTGGAATGCGTCATCCACCATTTTCATACACCGTATCAATTAATTTGAATACATGTG	360
Db	296977	TTTGAAAAAGGAGCATCTTCACATTTTCATATACACCGCTATATCATATGTTAAATTCATGAG	2970
Oy	361	TTAACCATTCACCTGGAAGATAAATAATC	387
Db	297037	TTAACCACTCCCGCAGATCAATAAATAC	297063
RESULT 8			
AE017028/c			
LOCUS	290117 bp	DNA	linear BCT 30-APR-2003
DEFINITION	Bacillus anthracis str. Ames section 5 of 18 of the complete genome.		
ACCESSION	AE017028	AE016879	
VERSION	AE017028.1	GI:30255149	
KEYWORDS	Bacillus anthracis str. Ames		
SOURCE	Bacteria; Filicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
ORGANISM	1 (bases 1 to 290117)		
REFERENCE	Read, T., Peterson, S., Tourasse, N., Bailly, L., Paulsen, I., Nelson, K., Mettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Oksstad, O., Helgason, E., Ralstone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., Deboy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, M., Peterson, J., Pop, M., Khouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaitte, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.		
AUTHORS	The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria		
JOURNAL	Nature	423 (6935), 81-86	(2003)
MEDLINE	22608414		
PUBMED	1271629		
REFERENCE	2 (bases 1 to 290117)		
AUTHORS	Read, T., Peterson, S., Tourasse, N., Bailly, L., Paulsen, I., Nelson, K., Mettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Oksstad, O., Helgason, E., Ralstone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., Deboy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, M., Peterson, J., Pop, M., Khouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaitte, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
FEATURES	Location/Qualifiers		
source	1..290117		
	/organism="Bacillus anthracis str. Ames"		
	/mol_type="genomic DNA"		
	/strain="Ames"		
	/db_xref="taxon:198094"		
	331..1929		
	/locus_tag="BA1197"		
	331..1929		
gene	/locus_tag="BA1197"		
CDS	/locus_tag="BA1197"		
	/note="similar to GP:10717143, and GP:10717143;		

identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product-"oligopeptide ABC transporter,
oligopeptide-binding protein"
/protein_id="AAP25159.1"
/db_xref="GI:30255150"
/translation-"MILGACSYQKDEPQANAKDGSKAKQVINTLVESEIPTMDTT
LASDASSKVMNNMTMEGLRLGKDKLVPVAKSHTSDEGKRYTEFLRDADKNSND
PVAKDFVYAMORAVNEDPTAAKSGYIMLDVKNAEKINKRELSPDLGVAKIDDTFEV
ELDNPVYFLDLIVYPLFPLNEKMEKQEGEGLTNTLLYNGPEVLNEMKHQDSQ
LKKNP5WMDQKEVLEEVNPNVYKDMAVNLNTSAIDRVGLSAELVYDKKGGDDK
TINDPYEFERFQKNPALKANKIRAKISSAYDRDGIYILNNGSGAIGLPSD
KGPDKODFERENKLSKVYDKEAKFEAGKKEGKDKITEFLNFDNESKITEYV
GKLEKMLPGLVITIKMOPPAOKLEKEDSGQFISFGMGDPDPDAITFLDMFTDGS
ONMKG5NACVYDGLIKKAKQEGSPKRMWMLDKAEKILFEDAAIAIVFORGRSILER
ETIKDYIHKYGEGLSPKMA5VEK"
complement(1956..2099)
/locus_tag="BA1198"
complement(1956..2099)
/locus_tag="BA1198"
/note-"identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product-"hypothetical protein"
/protein_id="AAP25160.1"
/db_xref="GI:30255151"
/translation-"MIGVITNSIGEKFLGGEKEMKFEVNSLVDVQSGWNEVGKGIYKK
RYY"
2731..3126
/locus_tag="BA1200"
2731..3126
/locus_tag="BA1200"
/note-"similar to SP:P32687, PID:409796, GB:U00096,
PID:1790459, SP:P32687, PID:409796, GB:U00096, and
PID:1790459; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product-"conserved hypothetical protein"
/protein_id="AAP25161.1"
/db_xref="GI:30255152"
/translation-"MWTLYSSPSCSTSCRAKMLLENHPIYTERNIPSDPLTEIKE
ILRMTEGSCDEITSTRSKVROELNVNLESLPLDDLYKMINDYPCOILLRPMIDEKRLQ
VGYNEDEIRRFRLRYVTFQDLREKQRLVN"
complement(3176..3850)
/locus_tag="BA1201"
complement(3176..3850)
/locus_tag="BA1201"
/note-"identified by match to Pfam protein family HMM
PF03772"
/codon_start=1
/transl_table=11
/product-"membrane protein, putative"
/protein_id="AAP25162.1"
/db_xref="GI:30255153"
/translation-"MDLEFLTSVLMYIGDIVLGGDNATVYALASRLPESKRNKAIL
IGTLALVLRILITLAVLYLDIPFLQIGGVLTLLAVNLITDNSMDLSSIGCKTTL
FOAVRTIVFADLVWGEDNVTAIGAAGHGRSLVIGLLISPIIIMSGSKLIMMERF
PFLICGAALITAYAGKMYTHEDRLATFFHNNSFTAGIPLYLFTLLICGIVQVRL
LRNWKH"
4112..4210
/locus_tag="BA1202"
4112..4210
/locus_tag="BA1202"
/note-"identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product-"hypothetical protein"
/protein_id="AAP25163.1"
/db_xref="GI:30255154"
/translation-"MICHSGESLHNNSKQEGRAVVEWILKELMIR"
4207..4860
/gene="mecc"
gene
CDS

CDS
/locus_tag="BA1203"
4207..4860
/gene="mecc"
/locus_tag="BA1203"
/note-"similar to GB:X03350, GB:X03350, GB:X03350,
GB:DD0137, GB:M21692, GB:M24317, GB:M24308, GB:M24309,
GB:M24310, GB:M24311, GB:M24312, GB:M24313, GB:M24314,
GB:M24315, GB:M24316, GB:M37067, GB:X15447, GB:X15448,
GB:X15449, GB:X15450, GB:X15451, GB:X15452, GB:X15453,
GB:X15454, GB:X15455, SP:P00325, SP:P00326, SP:P07327,
PID:1223729, PID:178098, PID:178100, PID:178111,
PID:219428, PID:28416, PID:598143, PID:825622, GB:X03350,
GB:X03350, GB:X03350, GB:DD0137, GB:M21692, GB:M24317,
GB:M24308, GB:M24310, GB:M24311, GB:M24312, GB:M24313,
GB:M24314, GB:M24315, GB:M24316, GB:M37067, GB:X15447,
GB:X15448, GB:X15449, GB:X15450, GB:X15451, GB:X15452,
GB:X15453, GB:X15454, GB:X15455, SP:P00325,
SP:P00326, SP:P07327, PID:1223729, PID:178098, PID:178100,
PID:178111, PID:219428, PID:28416, PID:598143, PID:825622,
GB:X03350, GB:X03350, GB:M24310, GB:M24311, GB:M24312,
GB:M24313, GB:M24314, GB:M24315, GB:M24316, GB:M37067,
GB:M24317, GB:M24308, GB:M24309, GB:M24310, GB:M24311,
GB:M24312, GB:M24313, GB:M24314, GB:M24315, GB:M24316,
GB:M37067, GB:X15447, GB:X15448, GB:X15"
/codon_start=1
/transl_table=11
/product-"negative regulator of competence Mecc"
/protein_id="AAP25164.1"
/db_xref="GI:30255155"
/translation-"MKPEITYIDIEDRGFNREIYDRESEELFEMEMDEARDHDF
FIDGRLIOVAOVADKGIETLVTKAELSDQOKLELPVGVKIIDIPDEGIESLFOOE
LVEVEEQAGTINREDCTPGFLIKFNDFEYISLHLLIEDIKEDISFEDRYVYV
EFDELVDHESEIDRLISTVLEYGESTLTTHRVSEKQIVKEHALETIRNFPKAK"
4933..6477
/gene="cls-2"
/locus_tag="BA1204"
4933..6477
/locus_tag="BA1204"
/note-"cls-2"
/locus_tag="BA1204"
/note-"similar to SP:066043, SP:066043, SP:066043,
SP:066043, SP:066043, and SP:066043; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product-"cardiolipin synthetase"
/protein_id="AAP25165.1"
/db_xref="GI:30255156"
/translation-"MKNTKLIFFVLLFALFVLSRMFIDVAFSDVIGIKOVSILGI
ISILFVTSAFILGCVIFLENRHPKSTLTVLIVGIFPVQPFVILFEGONFRKRMFO
KKALLDQAFLOYKRGHEDYERILRNKHOBELLFRLDRGALNISTQETRTLTNND
ETFOAILDGLKRAKNIHMEYIIVRDKLGTETKDILOKSGEVVRFVLDVNGSEK
LSKSYIELNDAGVEMIPPEVREPILDKLTGNRHKIVIDGNEGVGLNAGESEK
LGKDKYGFNRDTHLYLRGEAVOSLQILFIDFQVYMGGEAVLAPEYLOAKAVEGHWG
GVOIVAGGPNKMETIKHLYFAMTASARKIMVATPIFDDDLISLKLAAALAGIDV
RLMPSKPKDRIVYASRSYFPELLDGVIVIEKEKGLSKVIVYVSDIASLTAM
DMRSEFHNFEVNAFLVYDTSIRKLVQDFKDDLESSSIHVDRFKRLRHRIYESTYR
LLESLP"
6559..7803
/locus_tag="BA1205"
6559..7803
/locus_tag="BA1205"
/note-"identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product-"conserved hypothetical protein"
/protein_id="AAP25166.1"
/db_xref="GI:30255157"
/translation-"MIPPIFAKRENGKRIILLDHHDELLAHMKRRESFACAGKRV
LMRGKGRNHFARKVDSCLAFEAESTIHMVKGELLYTFWRQNRNVAITEHTLEPI
QOARDPIFERAGRIAYEQANLSIQKRYSVYQAOIOVIMIGQOKLKHSAV
MKFSSIMAFSLOSYPOLIFPSCSOSKMKCTFTFISTIAFSHIIILPIETTFE
EMLESVPQROREVLKREMKRRKDYFRONALDPIVNHKSLRLLYOOCKNPANPESPI
GVPLSPSFAQOTNSFTQWAFMKKICGELVAGVCCISIKYCYGTYKTKRRLPYPSQ
HMKVAVETMTFLCTAGVLLKYGMVYTKRIKRVAVPLKTGGEIILKHDEICLAAHLSLF

gene EAKFNMRGGKGDIIKTDREGIT"
7854. .9680
/gene="pepF-1"

gene 7854. .9680

```
/gene="pepf-1"
```

Query Match	54.48;	Score 210.6;	DB 1;	Length 290117;
Best Local Similarity	71.78;	Pred. No. 9.2e-38;		
Matches 276; Conservative	0;	Mismatches 109;	Indels 0;	Gaps 0;

[illegible]

RESULT 9	AE017002/c	AE017002	304680 bp	DNA	linear	BC1 16-MAY-2003
LOCUS						
DEFINITION		Bacillus cereus ATCC 14579	section 5 of 18	of the complete genome.		
ACCESSION		AE017002	AE016877			
VERSION		AE017002.1	GI:29894935			

Source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene
CDS

TITLE Direct Submission
JOURNAL Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Vilvert, Jouy en Josas 78352, France
FEATURES Location/Qualifiers

1. 304680
/organism="Bacillus cereus ATCC 14579"
/mol_type="genomic DNA"
/strain="ATCC 14579"
/db_xref="ATCC:14579"
/db_xref="taxon:226900"
68. 427
/locus_tag="BC1239"
68. 427
/locus_tag="BC1239"
/locus_tag="BC1239"
/codon_start=1
/transl_table=11
/product="hypothetical Membrane Spanning Protein"
/protein_id="AAO08224.1"
/db_xref="GI:29694936"
/translation="MARKRLKLYFLVPLVAMVLSMTKIELMALFLGVYEFISLNRI
QESLIEICFISFISIEFISVIVLNLNELFLRIHSFTLMKFWNIIIOAICTYIYFVL
KEIVGQVTFQDNKRME"
551. 2041
/locus_tag="BC1240"
551. 2041
/locus_tag="BC1240"
/codon_start=1
/transl_table=11
/product="L-lactate permease"
/protein_id="AAP08225.1"
/db_xref="GI:29694937"
/translation="MALLPIMMIFICLFKQKTSLSLSISYAVSGVILSLPTFQD
LGSVEVHATKGLGILFIVGVLEFGFLFLMNKKMKYIDQVAFLEEVNODRLQWML
IMQFGICPLIESGSGCIGMVAAPFELSGYKPPQAVILSGILASMSGMAVTGIT
IGSQLNMPLTTCITNTALISLTPVTRFAVYLSLVHVGQKQAVIEKKEGCFILSEL
GLYSANVAVSEIAGLISLSVITFTFGLLIKLKGSGQNLITTEPAATEKESVITKII
SPVIFLVYKLSLSQITKQWLPFAITTTMTFAISELMGASGMSHLATAGDTGTFEFV
FRPSMIKQSLISQITKQWLPFAITTTMTFAISELMGASGMSHLATAGDTGTFEFV
SVTFEIGALICGFLTSGNSNAHFIKLMQTAQNVALPMQVYTTIQNTASSAVTIACP
SRITLDAVLCNIPYRENEELKTKTLMIFGAVLLVVEVYIWMYLMRN"
2132. 2818
/locus_tag="BC1241"
2132. 2818
/locus_tag="BC1241"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08226.1"
/db_xref="GI:29694938"
/translation="MRLKNGVLKGVYIOSEIGKTPHYOVHLOGEGVDYRAIWNV
KQSGSYSEVYVFLNKKSAIHILPLPGFMETKNNEPKVALDYIRGNLPDSNOMI
PLPEKAGADNDLNKETIKRAIEKALITYAEGEKWGPENRPSDYPIHKPRNGIH
DIHMONGNEKMGQDGIWQDGLIHFEKEEWIGIFLAFQSSQCTDEGRHARVPV
EHMDYGNIK"
complement(2830. 3663)
/locus_tag="BC1242"
complement(2830. 3663)
/locus_tag="BC1242"
/codon_start=1
/transl_table=11
/product="Integral membrane protein"
/protein_id="AAP08227.1"
/db_xref="GI:29694939"
/translation="MELVNSIFQILLTSVQLSLISGVITVIGFLGLBSLRTMYNS
RAGRGKFLLTAAIGVDFELHGAHMKVLEPRHKIVAQFPPTDSQALGVQHYQYQW
KSYVORIGNFISGLTISQITLILMRFEVPESTYLFMTTEKTLASTSINIEMLQ
NMLSTLFTGIFGLNLPSEFMLEFLFAICISAHIALSKPIKSIDGVIVMFLV
LFLFNITITGFGYDSNQLIQVWKYNNYLAFSVALLFSCISITLVSEFSFYKIRGRSS
F"
3860. 4630
/locus_tag="BC1243"
3860. 4630
/locus_tag="BC1243"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08228.1"

gene
CDS
/db_xref="GI:29894940"
/translation="MKCPCHTGNNAEAKFCGNGCHSITEEVAWSGOEBEPDQAKA
RNEVEQAKRASGTFQFKEHALQAPAKMSAIEVRNIVSLVLIICFGACIFR
MMSASATYRFPADISTPTPEGESIVTFEFLIITLTFVGPILFEVSGMKMSFSFLE
VEGICGATPTIVILVLSFLFSFLIFLFLITLSVATYIGISITFAVAILKMGGLD
LVYTLIANVLGIAITFMVILMSYISTIIQALTKGITG"
4650..5885
/locus_tag="BC1244"
4650..5885
/locus_tag="BC1244"
/codon_start=1
/transl_table=11
/product="TG hypothetical 16680"
/protein_id="AAP08229.1"
/db_xref="GI:29894941"
/translation="MNWCTKCGAOFEDGVOPCONCGSKRGRPVYKKNMSGTKVITL
LALFVILIGILYGASYTVOAGVDNMTILIGRODGKLAETVADDPVMTRESIL
PLFSYIKENPSYVNELGKYLROGKORDGIERPDELITGDKYFEIREDRYKILAKYV
YTLTINKGSGYKMGNEKIDKTDKREKQYGFPLBNOYFOGDIKNDYKLSREK
VYLMQSOYNTVIDITLOGOYITVOITNAPGATLVNOKPYVALGEEITWGPVATDS
ATVYERNGESGRETTOVETITPLSSVLPPOKSTKTEVTYVNTPTPAGCHVYNGFI
PFDPSIRKLSADLVYLSKEBQKLARNETVYARHGIFQTKDMQAYEPAKOSYRBNPEP
KOTLDIEYVNEILKRSR"
complement(5931..6362)
/locus_tag="BC1245"
complement(5931..6362)
/locus_tag="BC1245"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08230.1"
/db_xref="GI:29894942"
/translation="WCAGLLAQLDPDEPEKRPAPVSLDQFSINATYISARAMLSN
DPTFQDIDMKATOFASASHACKSIFPVGSYKLIINFSDTITVTSPESDHALKIVPP
HSEALITGSPSHIQASTSGSSKVFLEFIHFSPDGPAPK"
5650..7894
/locus_tag="BC1246"
5650..7894
/locus_tag="BC1246"
/EC_number="1.6.99.3"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase"
/protein_id="AAP08231.1"
/db_xref="GI:29894943"
/translation="MNVYIIGDAAGSATQWQIVTNDENAVVYILEKEIYSYAQGL
PVYISGLASTEKLLAHNVKTFPRDKYGLDAVARYEATKVPDEKMYVAEPTKTFDE
PYDYLATGCVAPYPMPEGRLEOGYHLLKTIDARBIILKTLETPENVEPTIIGGA
IGLEMAEFVELGKKVRIERNDHIGTITYDADAEYIHKEDKHIIEILNENYKAK
GNERVEALETKRGTKRADLVASGVKPNDFLEGSTNIRKNHKAIEVNAVWQNVDD
VYAADDCAVTHVHVKIEHDIHPIGLTANRQGLAGLMLDKRRAFKGTIGTIGIIFND
LTGATVGNKEAGKGLHPIRYKTVGVDSYNNAGYYPNAKPLYKLILYSDFKQLGGGV
IGEEVDNRIVIVAMALEFNKMSIHLEDDVLSVAPYNSWMDPIQAAARAE"
8055..8465
/locus_tag="BC1247"
8055..8465
/locus_tag="BC1247"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08232.1"
/db_xref="GI:29894944"
/translation="MAINMKTIEMIAESNARHBEDEGHAVEEMKEEYCVGLDNATLYI
TKNVCFCFKVEVEFFQDHYVIGOEKDEYVEIKLKYDAITNSNLKTDNKNITLLEK
FANQGSITLSDINNYGTFKNMLFARQIKSIFKL"
complement(8496..10037)
/locus_tag="BC1248"
complement(8496..10037)
/locus_tag="BC1248"
/EC_number="1.6.99.3"
/codon_start=1
/transl_table=11
/product="NADH oxidase (NOXASE)"
gene
CDS

Query Match Best Local Similarity 71.4% Pred. No. 2,1e-37; Matches 275; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

gene
CDS

protein_id="AAP08233.1"
/db_xref="GI:29894945"
/translation="MQLNTFLHLHLEFSIDEIMPNHEVDKADLPYKLYKNVPTPLS
LEIPKLSNLSSTPTPLNEIGHYLMWSEGYTOLCOLMSGNRVLRSPSGGALPNELY
IYLKIDDDYPDGIYHYDAAHRLRLIEGNEEDSLDALGRNCILHPCGAAPVSTME
KNFEPYNNFSYRLQGDGVLIGOLLEKCKQPEYTGIVYPOFLDRNLHLIGISEGE
SVYATPLSTSEPEETNHFNDYHFNKRYKVTSHELLQOIPPLAHEHFFVSKVHDEKPIR
INEASIMESTAHFPOTLAHEEHYTOYHEHAVALPREVERSLIDFOLCKRRSPDQFTLT
KMDAEELATLIDLEASLISLPYSNDLQKYLENENRVSLLGYCYVNAKDIKNQAYVNSKT
HSIQPLRYGDLREYPLQSSMTMDNVNLEFVSLCHVVGKKQYRYNVALGRCYRIHOMKA
GLYHKLIVFAATAMGCMGHPLLSFDTPNSCDQLYGIDAGNETSLIQVPVAYRARNMLK
GALHNH"
complement(10053..11987)
/locus_tag="BC1249"
complement(10053..11987)
/locus_tag="BC1249"
/codon_start=1
/transl_table=11
/product="hypothetical Cytosolic Protein"
/protein_id="AAP08234.1"

Query Db 3 TCTTCTTCCTCTTCACCAAAATGAGAAACACTGCATCTAACAGTCACCAAGCCTAAA 62
||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244523 TCTTCTTCATCTTTGGAAAATAGTTGTGCAGCTTCGTCTAGTAAGTCACCTAACCTAAT 24446
||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Query Db 63 CCATGTGACCCGTGATATGCGGATACGGTTACCAACATCCCTAATATGAATATACATACAG 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244463 CCGTGTGACCGGAAATTTGGGAATGGTTGGCCAAATCCCTAATGCTAATAAGTCAATGAATG 24440

Query Db 123 TCTGTACCATTTCCATATTTACTACTTGTGTAAACCGCTAATACGACCGGTTTATAGAT 182
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244403 TCACGCGCATTTCCGGATTTGCTCTACTTTATTACTGCAAGTACCAACGGTTTATTGGAA 24434

Query Db 183 TTGTATAAATTTGAGCGACACTTTCATGCGTTTGTGCAATCTTCACGACGCTTAACC 242
||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244343 CGATATAAATTTTGGCACTCTTCATCATGAGCTGTAAACCGTCACGACCGCTTGTGC 24428

Query Db 243 ATAAATATTAATTAACATCCGCTTCATCTATAGCGGATTTGCGCGCGCTCAATTTGTGT 302
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244283 ATAAATATTAATTAACATCCGCTTCATCAATTTGCTACTTCAAGCTTTGACGAATTTGTGTC 24422

Query Db 303 TGAATGTGTCATCCAAATTCATCAATCAACCGCTGATCATATATATTTGAATCATGTGT 362
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244223 AAAAAGCGCTGCTCTCCCATATCAATTTCCGCCCGTATCATATATTTAAATTCATGATTT 24416

Query Db 363 AACCATTCACCTGAAGAATTAATAC 387
||||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244163 AACCATTCACCGCTATTAATAC 244139

RESULT 10
LOCUS AL596170 313450 bp DNA linear BCT 06-JUN-2002
DEFINITION *Listeria innocua* Cl1p1262 complete genome, segment 8/12.
ACCESSION AL596170
VERSION AL596170.1 GI:16414292
KEYWORDS
ORGANISM *Listeria innocua*
SOURCE *Listeria innocua*
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

REFERENCE 1
Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A.,
Baquero, F., Berche, P., Blocher, H., Brandt, P., Chakraborty, T.,
Charbit, A., Chetoui, F., Couve, E., de Daruvar, A., Delhoux, P.,
Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L.,
Dussanget, O., Entian, K.D., Fsihi, H., Portillo, F.G., Garrido, P.,
Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J.,
Jackson, D., Jones, L.M., Kaerst, U., Krell, J., Kuhn, M., Kunst, F.,
Kutkapat, G., Madueno, E., Maltournam, A., Vicente, J.M., Ng, E.,

LOCUS	AX413016	319630 bp	DNA	linear	PAT 02-SEP-2002
DEFINITION	Sequence 7 from Patent W00228891.				
ACCESSION	AX413016				
VERSION	AX413016.1	GI:21445474			
KEYWORDS					
SOURCE	<i>Listeria innocua</i>				
ORGANISM	<i>Listeria innocua</i>				
REFERENCE	1				
AUTHORS	Kunst, F. and Glaser, P.				
TITLE	<i>Listeria innocua</i> , genome and applications				
JOURNAL	Patent: WO 0228891-A 7 11-APR-2002; INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)				
FEATURES	Location/Qualifiers				
source	1. 319630 /organism="Listeria innocua" /mol_type="genomic DNA" /db_xref="taxon:1642"				
BASE COUNT	105207 a 55428 c 66726 g 92263 t			6 others	
ORIGIN					
Query Match	53.2%; Score 205.8; DB 6; Length 319630;				
Best Local Similarity	70.9%; Pred. No. 1.1e-36;				
Matches	273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;				
OY	3 TCTTCTCTCTCTTCACCAAAATGAGAACAACGTCATCTAACAGTGCACCAAGACTTAA 62				
Db	51123 TCCTCTCTCTCTTTTGGAAAATGAGACAGCAAGCATCAAGCATATACCAAGCCCTAGT 51064				
OY	63 CCATGTCACCCCTGATATGCGATACGGTTCACCAAACTCAATGAATGAAATCATACAG 122				
Db	51063 CCATGGAACCAAGAAATGGATAGCGCTCACCAACCAAGAGAAATGATATAAATC 51004				
OY	123 TCTGTATGCAATTCATATATCTACTGTTGTTAAACGCTAATACGACGCGTTTTAGAT 182				
Db	51003 TGTATCAGCGATTTCTGGGTATCTACTTATTAATGCGTAAACATTTGGTTATATAG 50944				
OY	183 TTGTATTAATAATTGAGCGACCATTTCAATCGCTTTGTGTCANCTCTTCACGACGTTAAC 242				
Db	50943 CGGTAAAGAAATTTTGTGCTACTGTTGTGTCATCGGTAAACCCCTTCACGACCATTTGTA 50884				
OY	243 ATTAATAATATATACATCCGCTTCATCTATGCGCATTTTCGCCGCGCTCAATTGTGT 302				
Db	50883 ATTAATAATATATAGCTGCTGCTTATCAATTTGCGATTTGCGCTTGTGCGGAATTTGCTGT 50824				
OY	303 TGAATAGTGTGCATCACCAATTTCAATACACGCTGATCAATATATTGAAATCATGTGTT 362				
Db	50823 AACAATGCTTCGTCGCGGAAGCATATATACCAACCTGTATCAATATGTTAAATTTCTTTCCA 50764				
OY	363 AACCATTTACCTGAGATATATATAC 387				
Db	50763 AGCCATTTCCGCTGAATATATATATGC 50739				
RESULT 12					
LOCUS	AL591981	347050 bp	DNA	linear	BCT 06-JUN-2002
DEFINITION	<i>Listeria monocytogenes</i> strain EGD, complete genome, segment 9/12.				
ACCESSION	AL591981	AL591824			
VERSION	AL591981.1	GI:16411141			
KEYWORDS					
SOURCE	<i>Listeria monocytogenes</i>				
ORGANISM	<i>Listeria monocytogenes</i>				
REFERENCE	1				
AUTHORS	Glaser, P., Frangoul, L., Buchrieser, C., Rusniok, C., Amend, A., Baqueiro, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T., Charlit, A., Chetouani, F., Couve, E., de Darvar, A., Deboux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurge, O., Entian, K.D., Fsih, H., Portillo, F.G., Garrido, P., Gautier, D., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, U., Jackson, D., Jones, L.M., Kaefer, U., Kreft, J., Kuhn, M., Kunst, F.,				

Kurupkat,G., Madueno,E., Maltournam,A., Vicente,J.M., Ng,E.,
Nedjari,H., Nordsiek,G., Novella,S., de Pablos,B., Perez-Diaz,J.C.,
Purcell,R., Remmel,B., Rose,M., Schluteter,T., Simoes,N.,
Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
Comparative genomics of *Listeria species*
Science 294 (3543), 849-852 (2001)

JOURNAL
MEDLINE
PUBMED
11679669
2 (bases 1 to 347050)

REFERENCE
AUTHORS
TITLE
JOURNAL
Glaser,P., Frangeul,L. and Rushtok,C.
Direct Submission
Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE
E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96. Fax: +33 (0)1 45 68 87 86.

COMMENT
FEATURES
source
Location/Qualifiers
1..347050
/organism="Listeria monocytogenes"
/mol_type="genomic DNA"
/strain="EGD-e"
/db_xref="taxon:1639"
complement(95..841)
/gene="lmo1688"
complement(95..841)
/gene="lmo1688"
/note="similar to glucose 1-dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="CAC99766.1"
/db_xref="GI:16411142"
/db_xref="SPTREMBL:O8Y6J6"
/translation="MNKVALYVSSRGIGREIALALAKRGDIANVSPNRKKAEPVO
OEIDOLGRCVIFKANGDVEKREYVDEFEGRDIFINNAAGVLRPLMELES
HMDWTMINAKALTRGDEPAKLMORHOSGKLTISUSISGIRYLEYTVGVSKAAVE
SLTRYLAELAFGLAVANVSGGLIETALNHPNKEELKDAVSKTPAGRMIERPDL
VNAVFLASEKMDIRGOTILVDGGRTLLV"
complement(join(845..1942,1948..1953))
/gene="lmo1689"
complement(845..1942)
/gene="lmo1689"
/note="similar to A/G-specific adenine glycosylase"
/codon_start=1
/transl_table=11
/protein_id="CAC99767.1"
/db_xref="GI:16411143"
/db_xref="SPTREMBL:O8Y6J5"
/translation="MRKKRLTWDETKITAFQEOALVSWYEAANKRVLPWRENTPEYRIW
VSEIMLOQTKVDVPIPFNREMTOPPTMESFVNADADILKAMEGLGYSRVRLQTA
MKQVADSGEVPDILTILSKGVPTAGALISLAYNQAEPAVDGVMRYIARLE
IGEDIMKASTRKIFEEVLXOLDKKNPAFNOGLMEIGALVCTPRKPMCLPDPFC
FAHKNGVETNPVKIKKVMKTKELSTIVISEDGKIAIEKRPENGLANMQFPPIE
ISKRENDEYAKLQFLHNTGLEVLLEDEPFAHKVSHLWKMDIRVALQSAIPNEN
WYFATEEEKKRLAFPPYOKMMQAWKDFEGE"
complement(850..855)
/gene="lmo1688"
complement(850..855)
/gene="lmo1688"
complement(1948..1953)
/gene="lmo1689"
join(2129..3123,3130..3147)
/gene="lmo1690"
2129..2134
/gene="lmo1690"
2143..3123
/gene="lmo1690"
/note="similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99768.1"
/db_xref="GI:16411144"
/db_xref="SPTREMBL:O8Y6J4"
/translation="MDTGHVVMGIALGALATVDPVVASOAAIGIMTATITIGSIP

DIDTVLKLNNADYIRNRGITHSLPMLAIWPLLSSILYLLFPATFYHLMLPFA
VGHILVDFENYGOAVPEPKEFWAFGEINTEDFIGSHVAIAAMLGSPVPT
FVTLVYLIAVYVAPROBMKHAVONLIPSEETITSTHFPQRYAVTQKHVY
VGRAPFRNISIEKPEPLRVPDNETIRSKKDKNLAIFSFVYWMRIEKKLDGTV
TFDLYRSNGHPVAVVVLDDDLKIVSSYGTWTFSTELKRLKLAHSVSI"
complement(join(3130..3147,3156..3617,3625..3630))
/gene="lmo1691"
3130..3147
/gene="lmo1690"
complement(3130..3147)
/gene="lmo1691"
complement(3156..3617)
/gene="lmo1691"
/note="similar to deoxyuridine triphosphate
nucleotidylhydrolases"
/codon_start=1
/transl_table=11
/protein_id="CAC99769.1"
/db_xref="GI:16411145"
/db_xref="SPTREMBL:O8Y6J3"
/translation="MKVRGFEVYNESKRFPPEOTISLPTRDGKSGAGYDFPSMETYTI
VPEKHIEFTDITKSTMOEDVELNITVRSSTIGIKGLLNGTGIIDSSYSSNGNDGN
IGIAIKNFSNEPVSIEAGERVAGVQKYLAVADTIVANESRVGVSIGR"
complement(3625..3630)
/gene="lmo1691"
complement(3629..3862)
/gene="lmo1692"
complement(3629..3862)
/gene="lmo1692"
/codon_start=1
/transl_table=11
/protein_id="CAC99770.1"
/db_xref="GI:16411146"
/db_xref="SPTREMBL:O8Y6J2"
/translation="MYTYLDELTAELMVKNPMLDKRQALMWIEMLSDESSYAKAGY
PYRGPEYADYVROQIERRHGSFLHYERKDPNK"
complement(join(3864..4673,4680..4685))
/gene="lmo1693"
complement(3864..4673)
/gene="lmo1693"
/note="similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99771.1"
/db_xref="GI:16411147"
/db_xref="SPTREMBL:O8Y6J1"
/translation="MKTISIVQOKKERYNIFIDEKYNVSDEVLARYOLMKGRVL
TEADIEIRQADMVRKGLKKAINFLSHRYRSEKIDYLRKQEMEPFADILKILAN
MYINDFEERALEYTKQITTKTPTREIREYKGLTREITSRYIEEVSDEQOLENA
TKOAKIMKRNKRSKMKMLQITIDLQKGFSTELAKAATSELDVADAEILQ
KQLEKTMKRNKRSISIAOKTITSLMOKGFSYDTIQSLTENEISFEERE"
complement(3872..3877)
/gene="lmo1692"
complement(3872..3877)
/gene="lmo1692"
complement(4680..4685)
/gene="lmo1693"
join(4757..5673,5681..5708)
/gene="lmo1694"
4757..4762
/gene="lmo1694"
4771..5673
/gene="lmo1694"
/note="similar to CDP-abequose synthase"
/codon_start=1
/transl_table=11
/protein_id="CAC99772.1"
/db_xref="GI:16411148"
/db_xref="SPTREMBL:O8Y6J0"
/translation="MHILTGTGFTIGDLVHELEKSDHELYITLRQKKNANWHYI
EVLNEDKLPNLEDLPVDCINLAGAGLMDKMTYDRKVIYNSRIEATSSALSTYKMK
KSKPKLIMASAIQATVSSKSTVYLDTEENTVADNPLGATVYEWKETSASDGLGIRV
VYARGLVLTGNGSFPVEKLFQYVTGFRGNGQWYSMIHVDVVAAILTIFDHEQ

REFERENCE	AUTHORS
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.	
1	Buchrieser, C., Frangeul, L., Couve, E., Rusnick, C., Esih, H., Deloux, P., Dussutge, O., Chetouani, F., Nedjari, H., Glaser, P., Kunst, F., Cozzari, P., Daniels, J., Goebel, W., Kreft, J., Kuhn, M., Ng, E., Vasquez-Boland, J., Dominguez-Bernal, G., Garrido-Garcia, P., Tierrez-Martinez, A., Amend, A., Chakraborty, T., Donnay, E., Hain, T., Berche, P., Charbil, A., Durant, L., Perez-Diaz, J.C., Baquero, F., Garcia del Portillo, F., Gomez-Lopez, N., Madueno, E., de Pablo, B., Weiland, J., Kaerst, U., Entian, K.D., Hauf, J., Rose, M., and Voss, H. Listeria monocytogenes genome, polypeptides and uses Patent: WO 0101118-A 2660 11-APR-2001;
TITLE	JOURNAL
PASTEUR (FR)	
FEATURES	SOURCE
Location/Qualifiers	
1..349980	/organism="Listeria monocytogenes"
/mol_type="genomic DNA"	/db_xref="taxon:1639"
/note="Seq 1 for big: original length: 2944528-splitlited into 10 <22>seq 0001 : From 0.000.001 to 0.349.980 length: 349.980 <22>seq 2855 : From 0.300.001 to 0.649.980 length: 349.980 <22>seq 2856 : From 0.600.001 to 0.949.980 length: 349.980 <22>seq 2857 : From 0.900.001 to 1.249.980 length: 349.980 <22>seq 2858 : From 1.200.001 to 1.549.980 length: 349.980 <22>seq 2859 : From 1.500.001 to 1.849.980 length: 349.980 <22>seq 2860 : From 1.800.001 to 2.149.980 length: 349.980 <22>seq 2861 : From 2.100.001 to 2.449.980 length: 349.980 <22>seq 2862 : From 2.400.001 to 2.749.980 length: 349.980 <22>seq 2863 : From 2.700.001 to 2.944.528 length: 244.528"	
BASE COUNT	99761 a 75437 c 61177 g 113605 t
ORIGIN	
Query Match	53.2% Score 205.8; DB 6; Length 349980;
Best Local Similarity	70.9% Pred. No. 1.le-36;
Matches	273; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
Db	3 TCCTTCCTCCTTCCACCAAAATGAGAACACTGCATCTACAGTCACCAGACCTAAA 62
211933	TCCTTCCTCCTTTGGAAAAATGAGCGCGAAGACGATCAGCATGTGCGCGACCCCTAGT 2119922
OY	63 CCAGTGACCGCTGATATGGATGCGGTTCACCAATCCTAATGAATAGAATCATACAG 122
211993	CCAGTGACCGAATAATGGATGAGCGGCTCACCAACCCAAAGAAATAAAGCATTAAT 2120522
OY	123 TCTGACCACTTTCATATTATCTACTTGTGTACCGGTAATACGACGGTTTTTAGAT 182
212053	TGATCTGCCATTTTCGGGTATCTACTTTATTATACGCTAAACAATTTGGTTTATTAGAC 212112
OY	183 TTGTATAAATTTGAGGACCACTTTCANCGCTTGTGTCAATCCTTCACGACGCTAAC 242
212113	CGGTAAAGAAATTTTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2121722
OY	243 ATAAAAAATTAATCACTCCGCTTCATCTATGGGATTTTCGCTGGCTTAATTTGTGT 302
212173	ATAAAAAATTAATCACTCGCTTCATCAATTTGGATTTTCGCTGCGGAAATTTGCTCT 2122322
OY	303 TGGAATGTGATCATCAATTCATACCACTGATATCAATTAATTTGAATCATGTGT 362
212233	AAAAATGTTTCATCGGAAAGATCATACCACTGATATCAATTAATTTGAATCATTTTCCA 2122922
OY	363 AACCAATTCACCTGAAGATTAATAC 387
Db	212293 AGCATTCGCTGAAATTAATATGC 212317
RESULT 15	
LOCUS	DEFINITION
BACSERA	24887 bp DNA linear BCF 09-DEC-1998
BACSERA/C	Bacillus subtilis phosphoglycerate dehydrogenase (sera), ypa,
	ferridoxin (fer), ypbB, resS, ypbD, ypbE, ypbF, ypgB, yphH,
	glutamate dehydrogenase (ypcA), ypdA, ypbB, ypbC, spore cortex

ACCESSION	VERSION	KEYWORDS
L47648	1	GI:1146195
SOURCE	ORGANISM	REFERENCE
Bacillus subtilis Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
AUTHORS	TITLE	JOURNAL
Roels,S., Driks,A. and Losick,R. J. Bacteriol. 174 (2), 575-585 (1992)	Characterization of spoIVA, a sporulation gene involved in coat morphogenesis in Bacillus subtilis.	Mol. Microbiol. 10 (2), 385-395 (1993)
REFERENCE	PUBMED	JOURNAL MEDLINE
2 (bases 1 to 24887) Sorokin,A., Zumstein,E., Azavedo,V., Ehrlich,S.D. and Serrin,P. The organization of the Bacillus subtilis 168 chromosome region between the spoVA and sera genetic loci, based on sequence data	92105027	Microbiol. 142 (Pt 8), 2005-2016 (1996)
AUTHORS	TITLE	JOURNAL MEDLINE PUBMED
3 (bases 1 to 24887) Sorokin,A., Azavedo,V., Zumstein,E., Galleron,N., Ehrlich,S.D. and Serrin,P. Sequence analysis of the Bacillus subtilis chromosome region between the sera and kdg loci cloned in a yeast artificial chromosome	79348929	Microbiology 142 (Pt 8), 2005-2016 (1996)
FEATURES	source	location/Qualifiers
1..24887 /organism="Bacillus subtilis" /mol_type="genomic DNA" /strain="Marburg" /sub_strain="168" /db_xref="taxon:1423" complement(1..41) /citation-[2] /citation-[3] /evidence=experimental complement(40..1617) /gene="serA" complement(40..1617) /gene="serA" /function="serine biosynthesis" /citation-[2] /citation-[3] /codon_start=1 /evidence=experimental /transl_table=1 /product="phosphoglycerate dehydrogenase" /protein_id="AAC83943.1" /db_xref="GI:1146196" /translation="MFRVLVSDKMSNDGLOPLIESDIEIVOKNVADAEDELTPDALLVSAATPTFEDLFNKMSLKTIGVRAGCVGNIDDEATKHGVIVINAPNGNTSTAHE TPNAISSLMHHTIOANISVKSRENNRPAVYGSSELYGTCTLTIVGLGRIGSETAORCAF GMYVHVDFPFLTERARKIGVNSRTPEEVLESADIITVHPILTKETGLINKETIAT KKGRLINCARGGIDEALLENCHVAGALDVVEBPVDNKLVDHPIIAIPH LGASTEAOLNAAOVAOSEVLEQFAKGLPVASAINLPMTADEPAKIKPYHQIGKIGS LVSGSEKIPQOVDAIOYEGLTAKLETSTFKALSGFLKRVDSVNEVNAAGAVER GISPEKSISSSEGQVDCISXVTGDSSTFTVATYIPHRGERTIVEINGENIDYPPG HLYTIQODTTVGIRYGRI LGNDINAITAMOVGRKEGGEAIMLSFDNHLEDKIVK ELINVDPDIVSKLIDLP"		
gene	CDS	gene
/gene="ypaA" 2114..2686 /gene="ypaA" /note="putative" /citation=[3]		

Db	21432	CGCTTCATCATTGGGATTTCAGCTTGCTGGCGAATCTGCCCTAAAAAGGCTCATCAC	21373
Oy	320	AATTCAATACCCTGTATCAATAATATGTGTTACCATTCACCTGAAGA	379
Db	21372	GATATCAATACCGCCGCTATCATCAATTAATAATCATTAATTCAGCCATTCAGCCGAGCT	21313
Oy	380	ATAAATAC	387
Db	21312	GTAATCC	21305

Search completed: September 12, 2003, 19:27:28
Job time : 1589.02 secs

THIS PAGE BLANK (USPTO)

C	9	676.8	51.9	319630	24	ABO67194	Listeria innocua C
C	10	676.8	51.9	3011208	24	ABO69245	Listeria innocua C
C	11	667.6	51.2	1960	22	AASB3683	S. epidermidis ope
C	12	600	46.0	1308	24	ABNB8457	Streptococcus poly
C	13	600	46.0	2155561	24	ABN71527	Streptococcus poly
C	14	599.4	45.9	1311	24	ABK75008	Bacillus lichenifo
C	15	582.4	44.6	2365589	24	ABA90521	Genomic sequence c
C	16	550.4	42.2	1308	25	ABX07474	S. pneumoniae type
C	17	550.4	42.2	1311	21	AAAS4516	ytic gene of Strept
C	18	550.4	42.2	1311	21	AAZS1826	Streptococcus pneu
C	19	550.4	42.2	5066	19	AAVS2212	Streptococcus pneu
C	20	550.4	42.2	2162598	25	ABS56454	Streptococcus pneu
C	21	549.6	42.1	1308	24	ABNB8458	Streptococcus poly
C	22	383.8	28.4	387	23	ABSA8886	Staphylococcus aur
C	23	367.2	28.1	372	23	AAS50706	Staphylococcus aur
C	24	336.8	25.8	2791	24	ABQJ0803	Listeria monocytog
C	25	318.4	24.4	738	23	AAS57835	Streptococcus pneu
C	26	308.6	23.6	675	23	AAS51384	Enterococcus faeca
C	27	308.6	23.6	678	23	AAS52955	Enterococcus faeca
C	28	286.4	22.7	298	23	AAS50205	Staphylococcus aur
C	29	286.4	22.7	298	23	AAS50723	Staphylococcus aur
C	30	293.6	22.5	580073	18	AATF8840	Mycoplasma genital
C	31	270.8	20.8	33140	22	ABA82787	Genomic fragment #
C	32	241.8	18.5	640681	24	ABA92787	Bacchara sp. genom
C	33	241.2	18.5	246	23	AAFS94379	Haemophilus influe
C	34	210.4	16.1	1512	22	AAF34379	Haemophilus influe
C	35	210.4	16.1	1512	24	ABK64943	DNA encoding Haemo
C	36	210.4	16.1	1515	23	AAS53235	Haemophilus influe
C	37	208.8	16.0	1830121	17	AAV42063	Haemophilus influe
C	38	203.4	15.6	78845	21	AAAF1608	N. meningitidis pa
C	39	203.4	15.6	349980	21	AAAF1608	N. meningitidis B
C	40	203.4	15.6	1437668	21	AAAR1480	N. meningitidis B
C	41	198.6	15.2	910715	20	AAAX0248	Borrelia burgdorfe
C	42	190.6	14.6	1500	25	ABZ593377	N. gonorrhoeae nuc
C	43	190	14.6	1512	23	AAS52517	E. coli DNA for ce
C	44	190	14.6	11574	22	AAAS6244	DNA encoding novel
C	45	178.6	13.7	1557	22	AAH66542	C glutamicum codin

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT;
 PI Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 DR P-PSDB: AAU33787.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS
 XX
 Claim 27; Seq ID No 4228; 511pp; English.
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/undepct_sequences.

Query Match	100.0%	Score 1305; DB 23;	Length 1305;
Best Local Similarity	100.0%;	Pred. No. 5.2e-290;	
Matches 1305; Conservative	0;	Mismatches	0; Gaps 0;

QY	1	ATGACTAACCCTATGTAGCTATTTAGT	AGGCGCTAATGTAGTAAATCTACAATTTT	60
Dp	1	ATGACTAACCCTATGTAGCTATTTAGT	AGGCGCTAATGTAGTAAATCTACAATTTT	60
QY	61	AATGAATAGTTGGAGACGTTCTTCGAT	GTGTGAAGATACGCCAGGTGTAACACAGAT	120
Dp	61	AATGAATAGTTGGAGACGTTCTTCGAT	GTGTGAAGATACGCCAGGTGTAACACAGAT	120
QY	121	CGTATTTATCTTCAGGTGAGTGTAA	CACATGATTCATATTATGTATCAGGTGT	180
Dp	121	CGTATTTATCTTCAGGTGAGTGTAA	CACATGATTCATATTATGTATCAGGTGT	180
QY	181	ATTGAATTTGGTGATGCACCATTTCC	AAACACAAATTTAGAGCGCAGACAGAAATGCCGAT	240
Dp	181	ATTGAATTTGGTGATGCACCATTTCC	AAACACAAATTTAGAGCGCAGACAGAAATGCCGAT	240
QY	241	GATGAGCGGAGTATATTTTATTTAT	GTGTAACTGTCGCGGAAGGATTGACACAAAGCAT	300
Dp	241	GATGAGCGGAGTATATTTTATTTAT	GTGTAACTGTCGCGGAAGGATTGACACAAAGCAT	300
QY	301	GAATGGTCGCTCAAAATTTTATACAA	ATCTAAAAAACCGGTGATTACGGTTAACAA	360
Dp	301	GAATGGTCGCTCAAAATTTTATACAA	ATCTAAAAAACCGGTGATTACGGTTAACAA	360
QY	361	GTAGATAATATGGAATCGCTACAGAC	GTGTATGATTTCTATTCTATTAGATTGGTGAA	420
Dp	361	GTAGATAATATGGAATCGCTACAGAC	GTGTATGATTTCTATTCTATTAGATTGGTGAA	420
QY	421	CCGATATCCAAATATCAGGCTCACAT	GGTTAGTCTTGGTCTGTAGATGCAGTGTGTT	480
Dp	421	CCGATATCCAAATATCAGGCTCACAT	GGTTAGTCTTGGTCTGTAGATGCAGTGTGTT	480
QY	481	TTCATATTTGGTGAGAGGAGAAACAT	CTTATGATGAAGATACAAATTGCATATCCATT	540
Dp	481	TTCATATTTGGTGAGAGGAGAAACAT	CTTATGATGAAGATACAAATTGCATATCCATT	540

QY	541	TTTTGAGACACCAACGTAAGTAATCAAGTTTACTAAATGCTATTTTAAAGTGAAGATCGC	600
Db	541	ATTTGAGACACCAACGTAAGTAATCAAGTTTACTAAATGCTATTTTAAAGTGAAGATCGC	600
QY	601	GTTATTCGTTTCTAAATGTTGCAGGAGCAACGAGACCGCTATTGCATACAGATATAGTTAT	660
Db	601	GTTATTCGTTTCTAAATGTTGCAGGAGCAACGAGACCGCTATTGCATACAGATATAGTTAT	660
QY	661	GATGAGACAAAGTATATGTTTTTAATGATATCTGCTGTAATGCGTAAAAAGGAAAAATATAT	720
Db	661	GATGAGACAAAGTATATGTTTTTAATGATATCTGCTGTAATGCGTAAAAAGGAAAAATATAT	720
QY	721	GAATCAACTGTGAAGAAATATTCAGTATTTAAAGACCTTTAAAGCGATTTGAACCTTCAAAATGT	780
Db	721	GAATCAACTGTGAAGAAATATTCAGTATTTAAAGACCTTTAAAGCGATTTGAACCTTCAAAATGT	780
QY	781	GTTTTAGTGGTATTTAGTATGACGAACAAGGCATTTATGACAAAGATTAACGATGTTGCAGGA	840
Db	781	GTTTTAGTGGTATTTAGTATGACGAACAAGGCATTTATGACAAAGATTTAAAGCGATTTGAACCTTCAAAATGT	840
QY	841	TATGCACATGAAACAAGGTAAAGACAGTCGTCGTAATTTGCTAAATAAATGGAATAGTCTGAGAA	900
Db	841	TATGCACATGAAACAAGGTAAAGACAGTCGTCGTAATTTGCTAAATAAATGGAATAGTCTGAGAA	900
QY	901	AAAGATAGTAAACAGATGAGAAATAATTTGAGATCAAGTACGTAAAGAAATTCCAATTTT	960
Db	901	AAAGATAGTAAACAGATGAGAAATAATTTGAGATCAAGTACGTAAAGAAATTCCAATTTT	960
QY	961	GATTATGACACAATTTGCTTTGTCGTCTGCTAAAGACGCAACAGATTAAGCTATATTC	1020
Db	961	GATTATGACACAATTTGCTTTGTCGTCTGCTAAAGACGCAACAGATTAAGCTATATTC	1020
QY	1021	CCTTACATTAATGAGCAAGTGAACCATTAACCAACGTTGCTCAAACTTCAACTTTAAAT	1080
Db	1021	CCTTACATTAATGAGCAAGTGAACCATTAACCAACGTTGCTCAAACTTCAACTTTAAAT	1080
QY	1081	GAAATCGTACTGATGCAATTTCCATGATACACCCTTACACCAACGCAAAAGGTAGACGTTTG	1140
Db	1081	GAAATCGTACTGATGCAATTTCCATGATACACCCTTACACCAACGCAAAAGGTAGACGTTTG	1140
QY	1141	AATGCTCTTTATGGAACGCAAGTTGCTATAGAACCCGACGACATTTGTTGATTTTGTAAAT	1200
Db	1141	AATGCTCTTTATGGAACGCAAGTTGCTATAGAACCCGACGACATTTGTTGATTTTGTAAAT	1200
QY	1201	GATGTAGAAATTAATGCAATTTTCTTATTAACGCTATTTTAAGAAATCAAAATCCGTGCGCT	1260
Db	1201	GATGTAGAAATTAATGCAATTTTCTTATTAACGCTATTTTAAGAAATCAAAATCCGTGCGCT	1260
QY	1261	TTTGCGTTTGAAGGTACACCAATTCATATATATATAGCTGGAAGAGCA 1305	
Db	1261	TTTGCGTTTGAAGGTACACCAATTCATATATATATATAGCTGGAAGAGCA 1305	
RESULT 2			
ID	AA534865	standard; DNA; 1311 BP.	
AC	AA534865;		
XX	13-FEB-2002	(first entry)	
DE	Staphylococcus aureus DNA for cellular proliferation protein #1177.		
KW	Antisense; ds: prokaryotic cellular proliferation gene;		
XX	antibiotic; antibacterial; drug design.		
OS	Staphylococcus aureus.		
PN	MO200170955-A2.		
XX	27-SEP-2001.		
PD	21-MAR-2001; 2001MO-US09180.		
PF			

XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR P-PSDB: AUN37006.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 8502; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from Wipo at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1311 BP; 451 A; 181 C; 277 G; 402 T; 0 other;
Query Match 100.0%; Score 1305; DB 23; Length 1311;
Best Local Similarity 100.0%; Pred. No. 5.2e-290;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACTAAACCTATAGCTATTTAGGTAGGCGCTTAATGTAGTAATCTACAAATTTT 60
DB 1 ATGACTAAACCTATAGCTATTTAGGTAGGCGCTTAATGTAGTAATCTACAAATTTT 60
QY 61 AATAGATAGTTGGAGACGTTGTTGATTGGAAGATACGCCAGGTGTACACAGAT 120
DB 61 AATAGATAGTTGGAGACGTTGTTGATTGGAAGATACGCCAGGTGTACACAGAT 120
QY 121 CGATTTATTTCTTCAGGTAGTGGTTAACACATGATTCAATTTAATTTGTTACAGGTGG 180
DB 121 CGATTTATTTCTTCAGGTAGTGGTTAACACATGATTCAATTTAATTTGTTACAGGTGG 180
QY 181 ATTGAATTTGGTATGATGACCATTTCCAAACAAATTAGAGCGGACAGAAATCGCCATA 240
DB 181 ATTGAATTTGGTATGATGACCATTTCCAAACAAATTAGAGCGGACAGAAATCGCCATA 240
QY 241 GATGAAGCGGATGTTATTATTTTATGTTAACTGCGTGAAGGATGTACACAAACGAT 300
DB 241 GATGAAGCGGATGTTATTATTTTATGTTAACTGCGTGAAGGATGTACACAAACGAT 300
QY 301 GAAATGCGTCGTCAAATTTTATACAAATCTAAAAAACCCTGTTATTAGCGGTTAACAAA 360
DB 301 GAAATGCGTCGTCAAATTTTATACAAATCTAAAAAACCCTGTTATTAGCGGTTAACAAA 360
QY 361 GTAGATAAATATGAAAATGCGTACAGACGTGTATGATTTCTATTCAATTAGATTTGGTGA 420

DB 361 GTAGATAAATATGAAAATGCGTACAGACGTGTATGATTTCTATTCAATTAGATTTGGTGA 420
QY 421 CCGTATCCAAATATACAGGTCACATGCTTTAGGTCCTGGTACATGTTAGATGAGTTGGT 480
DB 421 CCGTATCCAAATATACAGGTCACATGCTTTAGGTCCTGGTACATGTTAGATGAGTTGGT 480
QY 481 TCTCATTTTGGTGAAGAGAGAGAATCCTTATGATGAAGATATCAATTCGATATCCATT 540
DB 481 TCTCATTTTGGTGAAGAGAGAGAATCCTTATGATGAAGATATCAATTCGATATCCATT 540
QY 541 ATTGACGACCAAAACGTAGGTAATCAAGTTTATAGTAATGTAATTTAGGTGAAGATCCG 600
DB 541 ATTGACGACCAAAACGTAGGTAATCAAGTTTATAGTAATGTAATTTAGGTGAAGATCCG 600
QY 601 GTTATGCTTTCTAATGTTGACGAGGACACAGAGACGCTATTGATACAGATTAATTA 660
DB 601 GTTATGCTTTCTAATGTTGACGAGGACACAGAGACGCTATTGATACAGATTAATTA 660
QY 661 GATGACCAAGATTAATGTTTAAATGATACGCTGTATGCGTAAAAAAGGAAAGATATAT 720
DB 661 GATGACCAAGATTAATGTTTAAATGATACGCTGTATGCGTAAAAAAGGAAAGATATAT 720
QY 721 GAATCAACGTGAGAAATATTCAGTATTAAAGCTTTAAAGCGATTGAACGTTCAAAATGTT 780
DB 721 GAATCAACGTGAGAAATATTCAGTATTAAAGCTTTAAAGCGATTGAACGTTCAAAATGTT 780
QY 781 GTTTTATGCTGTTATTTATGATGACAGAAACAGGATTTATGACAGATTAACGTTGACGA 840
DB 781 GTTTTATGCTGTTATTTATGATGACAGAAACAGGATTTATGACAGATTAACGTTGACGA 840
QY 841 TATGACATGACAAAGGTAAGGACGTCGATTTGATTAATTAATGAGATTAAGTGTGGA 900
DB 841 TATGACATGACAAAGGTAAGGACGTCGATTTGATTAATTAATGAGATTAAGTGTGGA 900
QY 901 AAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTAAAGATTAAGTAAATTTTAA 960
DB 901 AAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTAAAGATTAAGTAAATTTTAA 960
QY 961 GATTATGACAAATTTGCTTTGCTGCTGAAGAACGCAAGATTAAGTAAATTTTAA 1020
DB 961 GATTATGACAAATTTGCTTTGCTGCTGAAGAACGCAAGATTAAGTAAATTTTAA 1020
QY 1021 CCTTACATTAATGATGACAAAGTAAACCATTAATTAATTAATTAATTAATTAAT 1080
DB 1021 CCTTACATTAATGATGACAAAGTAAACCATTAATTAATTAATTAATTAATTAAT 1080
QY 1081 GAACTGTTACTGATGCAATTTTCCATGAACCTTACACCAACAGACAAAGGTAGACGTTTG 1140
DB 1081 GAACTGTTACTGATGCAATTTTCCATGAACCTTACACCAACAGACAAAGGTAGACGTTTG 1140
QY 1141 AATGCTTTTATGACACAGTTCGATATGAACCAACGCAATTTGTTGATTTGTTAAT 1200
DB 1141 AATGCTTTTATGACACAGTTCGATATGAACCAACGCAATTTGTTGATTTGTTAAT 1200
QY 1201 GATGTGAATTAATGATTTTCTTAAACGCTATTAGAGATCAAAATCCGTGCCGT 1260
DB 1201 GATGTGAATTAATGATTTTCTTAAACGCTATTAGAGATCAAAATCCGTGCCGT 1260
QY 1261 TTTGGTTTGAAGGTACACCAATTCATTTATATAGCTGGAAGAGA 1305
DB 1261 TTTGGTTTGAAGGTACACCAATTCATTTATATAGCTGGAAGAGA 1305
RESULT 3
AAF86461 standard; DNA: 1311 BP.
ID AAF86461.
XX
AC AAF86461.
XX
DT 26-JUN-2001 (first entry)
XX
DE Staphylococcus aureus typhc coding sequence.

QY 1261 TTGTGTTTGAAGTACCAATTCATATTATAGCTCGAAGAAGA 1305
|||||
DB 1261 TTGTGTTTGAAGTACCAATTCATATTATAGCTCGAAGAAGA 1305

RESULT 5
AAV74669
ID AAV74669 standard; DNA; 3621 BP.

XX AAV74669;
XX
DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contlg SEQ ID #358.
XX
XX Computer readable medium; vaccine; S. aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
XX Staphylococcus aureus.
OS
FH Key Location/Qualifiers
FT misc_feature 481..540
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc_feature 2281..2340
FT /tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

XX EP786519-A2.
XX
PD 30-JUL-1997.
XX
XX
XX 07-JAN-1997; 97EP-0100117.
XX
XX
XX 05-JAN-1996; 96US-0009861.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX Rosen CA;
XX
XX WPI; 1997-374922/35.
XX
XX
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus
XX stored on computer readable medium and used in the production of
XX anti-S. aureus vaccines
XX
XX
XX Claim 1: Page 1241-1243; 3271pp; English.

CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S. aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S. aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S. aureus in a sample. S. aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences

CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S. aureus DNA sequences contained on the
CC computer readable medium.
XX
XX Sequence 3621 BP; 1279 A; 442 C; 692 G; 1085 T; 123 other;

Query Match 94.3%; Score 1230.2; DB 18; Length 3621;
Best Local Similarity 94.6%; Pred. No. 1e-272;
Matches 1235; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGCTATTTAGGTAGGCGCTAATGATGTAATCTCAATTTT 60
|||||
DB 2130 ATGACTAAACCTATAGCTATTTAGGTAGGCGCTAATGATGTAATCTCAATTTT 2189

QY 61 AATGAGATAGTGGAGAACGCTGTTTCATTTGTTGAAGATACGACAGTGAACAGAGAT 120
|||||
DB 2190 AATGAGATAGTGGAGAACGCTGTTTCATTTGTTGAAGATACGACAGTGAACAGAGAT 2249

QY 121 CGTATTTATCTTCACGAGTGTGTTTAAACATGATTTCAATATTATTGATACAGTGGT 180
|||||
DB 2250 CGTATTTATCTTCACGAGTGTGTTTAAACATGATTTCAATATTATTGATACAGTGGT 2309

QY 181 ATTGAAATTGGTGATGCACCATTTCCAAACACAATATAGAGCGAGCAGAAATCGGCAT 240
|||||
DB 2310 NNN 2369

QY 241 GATGAAGCGGATGTTATTTATTTATGTTAACTGTCGTTGAAGATTTGACACAAAGCAT 300
|||||
DB 2370 GATGAAGCGGATGTTATTTATTTATGTTAACTGTCGTTGAAGATTTGACACAAAGCAT 2429

QY 301 GAAATGCTGCTCAATTTTATACAAATCTAAAAAACCCTGATTTACCGTTAACAA 360
|||||
DB 2430 GAAATGCTGCTCAATTTTATACAAATCTAAAAAACCCTGATTTACCGTTAACAA 2489

QY 361 GTAGATTAATATGAAATGCGTACAGACGTTATGATTTGATTTAGATTTGGTGA 420
|||||
DB 2490 GTAGATTAATATGAAATGCGTACAGACGTTATGATTTGATTTAGATTTGGTGA 2549

QY 421 CCGTATCCAATATCAGAGTGCACATGTTAGTCTTGATCTGTTAGATCAGTTGTT 480
|||||
DB 2550 CCGTATCCAATATCAGAGTGCACATGTTAGTCTTGATCTGTTAGATCAGTTGTT 2609

QY 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAAGATACATTTGACTATTCATT 540
|||||
DB 2610 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGAAGATACATTTGACTATTCATT 2669

QY 541 ATTGAGACGACCAAGCTAGGTAAATCAGTTAATGATTTAGTGAAGATGCC 600
|||||
DB 2670 ATTGAGACGACCAAGCTAGGTAAATCAGTTAATGATTTAGTGAAGATGCC 2729

QY 601 GTTATCGTTTCTAATGTTCCAGGACACAGAGACGCTATTGATACAGATATAGTTAT 660
|||||
DB 2730 GTTATCGTTTCTAATGTTCCAGGACACAGAGACGCTATTGATACAGATATAGTTAT 2789

QY 661 GATGACAAAGATTATGTTTATTCATCTGCTGGTATGCTGTTAAAAAAGAAAGATAT 720
|||||
DB 2790 GATGACAAAGATTATGTTTATTCATCTGCTGGTATGCTGTTAAAAAAGAAAGATAT 2849

QY 721 GAATCAACGAGAAAATATTCAGTATTAAGAGCTTTAAAGCGATTGAAGCTTCAATGTT 780
|||||
DB 2850 GAATCAACGAGAAAATATTCAGTATTAAGAGCTTTAAAGCGATTGAAGCTTCAATGTT 2909

QY 781 GTTTAGTGTATTTATGATCAGAACAGGCAATTTATTTGAACAAGATTAACGTTGACAGA 840
|||||
DB 2910 GTTTAGTGTATTTATGATCAGAACAGGCAATTTATTTGAACAAGATTAACGTTGACAGA 2969

QY 841 TATGCAATGAACAGATTAAGAGATGCTGATTTGTTAAATTAATGGATCTGTGGAA 900
|||||
DB 2970 TATGCAATGAACAGATTAAGAGATGCTGATTTGTTAAATTAATGGATCTGTGGAA 3029

QY 901 AAAAGATAGTAAACGATGAAGAAATTTGAAGATGATGCTGAAGAAATTTCAATTTTA 960
|||||
DB 3030 AAAAGATAGTAAACGATGAAGAAATTTGAAGATGATGCTGAAGAAATTTCAATTTTA 3089

OY	961	GATTATGCACAAATTTGCTTTGTGTGTGCTTAAGACGCACAAATTTACGTACATTATTC	1020
Db	3090	GATTATGCACAAATTTGCTTTGTGTGTGCTTAAGACGCACAAATTTACGTACATTATTC	3149
OY	1021	CCTTACATTAATTAAGCAAGTGAACCATTAATAAAGCTGTCAAAGTCAACTTAAT	1088
Db	3150	CCTTACATTAATTAAGCAAGTGAACCATTAATAAAGCTGTCAAAGTCAACTTAAT	3209
OY	1081	GAAGTCGTACTGATGCATTTCCATGAAACCTTACACACAGACAAAGTAGACGTTTG	1140
Db	3210	GAAGTCGTACTGATGCATTTCCATGAAACCTTACACACAGACAAAGTAGACGTTTG	3265
OY	1141	AATGTCCTTTTATGCAACAAGTTCCTATTAGAACACCGACATTTGTTGTAATTTGTTAAT	1200
Db	3270	AATGTCCTTTTATGCAACAAGTTCCTATTAGAACACCGACATTTGTTGTAATTTGTTAAT	3329
OY	1201	GATGTAGCAATTAATGCATTTTCTTATTAACGCTATTATTAGAAATCAAAATCCGTCCGCT	1260
Db	3330	GATGTAGCAATTAATGCATTTTCTTATTAACGCTATTATTAGAAATCAAAATCCGTCCGCT	3389
OY	1261	TTTGCGTTTGAAGTAGACCAATTCATATTATTACCTGCAAGAAGA	1305
Db	3390	TTTGCGTTTGAAGTAGACCAATTCATATTATTACCTGCAAGAAGA	3434

XX	Sequence	1332 BP, 465 A, 190 C, 267 G, 410 T; 0 other:
SQL	70.6%; Score 921; DB 24; Length 1332;	
	Best Local Similarity 81.6%; Pred. No. 7.7e-202;	
	Matches 1065; Conservative 0; Mismatches 240; Indels 0; Gaps 0;	
QY	1 ATGACTAAACCTATAGTACGTAATGTTAGGTAGGCGCTAAATGTAAGTAAATCTACAAATTTT	60
DB		
QY	22 ATGACTAAACCGATAGTACCAATGTTAGGAAACCCAAATGTAGAAATCTACAAATTTT	81
DB		
QY	61 AATGAAATAGTTGGGAACGTTTCCATTGTTGAAGTACGCCAGGTGTAAACGAAAT	120
DB		
QY	82 AATAAGATGTCGGGGAACGTTACTATCTAGTAAAGATACGCCAGGCGTTACTGTAAT	141
DB		
QY	121 CGTATTTATTTCTTACGTTAGTGGTTAACACATGATTTTCAATATATGATACAGTGGT	180
DB		
QY	142 CGTATTTATTTCTTACGTTAGTGGTTAACACATGATTTTCAATATATGATACAGTGGT	201
DB		
QY	181 ATGAAATTTGATGATGACCAATTCACAAACACAAATTTAGAGCGGACGAGAAATCGCAT	240
DB		
QY	202 ATGAAATTTGATGATGACCAATTCACAAACACAAATTCAGAGCGGACGAGAAATCGCAT	261
DB		
QY	241 GATGAAGCGGATGTTATTTATTTTATGTTTAACTGCGCTGAAGATTTACACAAAGCAT	300
DB		
QY	262 GAAAGAGAGATGTCATCATTTTATGTCATATGTAAAGAGAGCTTACACAAAGTAC	321
DB		
QY	301 GAAATGGTCGTCACAAATTTTATACAAATCTAAACAAACCGGTCGTTATAGCGTTAACAA	360
DB		
QY	322 GAAATGGTCGTCACAAATCTTATTAATCTAAGAAACCTGTTGTAATGCTGTCAATAA	381
DB		
QY	361 GTAGATTAATATGGAATCGTACACAGCTGATGATTTCTATTCATTAAGATTGGTGAA	420
DB		
QY	382 GTGATTAATCTTGAATCGTAAATGATATGATGATTTCTATTCATTAAGATTGGTGAA	441
DB		
QY	421 CCGTATCCAATATCAGGTCACATGTTAGTCTTGGTGACTGTTAGATGCACTGTT	480
DB		
QY	442 CCATATCTTATTTCTGTTTACATGATGATGATGATGATGATGATGATGATGATGATGAT	501
DB		
QY	481 TCTCATTTTGGGAGAGAGAGAGATGCTTATGATGATGATGATGATGATGATGATGAT	540
DB		
QY	502 GAAATCTTAAATTAAGAAATCAGAAATCCTTATGAGAGATGATGATGATGATGATGAT	561
DB		
QY	541 ATTGACGACCAACAAAGTATGATTAATCAAGTTAGTAAATGCTATTTTATGAGTGAAGATGC	600
DB		
QY	562 ATCGTATACCTTAATGTTGTTAAATCTAGCTTGGTCAATGCTATTTTATGAGTGAAGATGC	621
DB		
QY	601 GTTATCGTTTCTAATGTTTGAAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	660
DB		
QY	622 GTTATCGTTTCTAATGTTTGAAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	681
DB		
QY	661 GATGACCAAGATTAATGTTTGAAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720
DB		
QY	682 GATGACCAAGATTAATGTTTGAAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	741
DB		
QY	721 GAATCAATCGAATAATTTATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	780
DB		
QY	742 GAATCAATCGAATAATTTATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	801
DB		
QY	781 GTTATCGTTTCTAATGTTTGAAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840
DB		
QY	802 GTTATCGTTTCTAATGTTTGAAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	861
DB		
QY	841 TATGACCAATGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	900
DB		
QY	862 TATGACCAATGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	921
DB		
QY	901 AAAGTAACTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	960
DB		
QY	922 AAAGTAACTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	981
DB		
QY	961 GATATGCAAAATGCTTTTGTGCTCTCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1020
DB		

Db 2065970 TATGCGCATATGCGGACCGCGCATATATGTAGTGAACAATGGATGCAATTAAC 2065911
Qy 901 AAGATAGTAAACGATGAAGAAATTTGAAGATACAGTACAGTAAAGATTCATTTTAA 960
Db 2065910 AAGATGAAAAAACATTAATGATGACAGAGATTTGGGAGCGATTCATTTCTTA 2065851
Qy 961 GATTATGACAAATTTGCTTTGTGTGCTGAAGAACGACAAAGATTAAGTATATTC 1020
Db 2065850 AGCATGACCAATTTGTTGCTGATGCTAAACAAACAGCTTAACCACTTATTC 2065791
Qy 1021 CCTTACTTATGAGCAAGTGAAGAACCATTAACATTTGTTCAAGTTCACTTAAAT 1080
Db 2065790 CCGTCATTTAAACAGTAAGTACGATTAATTCATTCAGCGCATATAGTATCTAAT 2065731
Qy 1081 GAAGTCCTTACTGATGCAATTTCCATGACCCCTACACCAAGAGTAGAGTGTG 1140
Db 2065730 GATGTTTATGATGATGCGGTTGCAATGATCATTGATTAAGTAAAGCACTT 2065671
Qy 1141 AATGCTTTTATGACACAAAGTTGCTATAGAACCGACGATTTGTTATTTTAT 1200
Db 2065670 AATATTTCTATACAAACAAAGTGTGTTAAACCGCACATTTGTTATTTTAT 2065611
Qy 1201 GATGTAATTAATGCAATTTTCTTATAACGCTATTTAGAAATCAATCCGTGCGCT 1260
Db 2065610 GATTCAGAACTAAATGATTTCTTATGAACGTTTCTTAGAAAACGATTTAGAGAA 2065551
Qy 1261 TTTGGTTTGAAGTACCAATTCATATTATACCTGGAAGAG 1304
Db 2065550 TTTCCGTTTATGATGATGACCAATTCGATTAATTCCTGTAAGCG 2065507

RESULT 11

AAH53683 standard; DNA; 960 BP.

XX ID AAH53683 standard; DNA; 960 BP.
XX AC AAH53683;
XX DT 03-SEP-2001 (first entry)
XX S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2759.
XX DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2759.
XX KM Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX KM vaccination; endocarditis; ds.
XX OS Staphylococcus epidermidis.
XX PN W0200134809-A2.
XX PD 17-MAY-2001.
XX PF 09-NOV-2000; 2000WO-US30782.
XX PR 09-NOV-1999; 99US-0164258.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Kimmberly MJ;
XX DR WPI: 2001-316495/33.
XX DR P-PSDB; AAG82833.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX PT useful for vaccinating against infections, e.g. endocarditis -
XX PS Claim 8; Page 718-719; 2188pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
XX CC (I) and (II) can have antibacterial activity and therefore can be used
XX CC in vaccination. The nucleic acids (I) may be used to produce the
XX CC S. epidermidis polypeptides (II) via the production of vectors
XX CC containing them which are used to produce hosts cells which express the

CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
XX

Sequence 960 BP; 336 A; 131 C; 189 G; 304 T; 0 other;

Query Match 51.2%; Score 667.6; DB 22; Length 960;
Best Local Similarity 81.2%; Pred. No. 1.e-143;
Matches 775; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 352 GTTAAACAAGTAGATTAATGGAATGCGTAGACAGCGTATGATTTCTATTCATTTAGGA 411
Db 1 GTGATTAAGTTGATTAATCTTGAAATCGTATGATATCTATGATTTCTATTTTGGC 60
Qy 412 TTTGCTGAACCGTATTCGAATATCAGGCTCAATGTTAGGCTTGGTGAATTTAGAT 471
Db 61 TTTGAGAGTCATATTCCTATTTTGTGTTACATGATTTAGACTTTGGAGATTTCTGAT 120
Qy 472 GCAATGTTTCTCATTTTGGTGAAGAGAGAAAGATCTTATGATGAAGATACAAATCGA 531
Db 121 GCAGTTGTGAAACCTTTAATTAAGAAATCGAAGATCTTATGACGATGATGATGAT 180
Qy 532 CTATCCATTAATTTGACGACCAACAGTGAATCAATCAATTTAGTAATTTAGT 591
Db 181 CTTTCTATCATCGGTATGACCTTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240
Qy 592 GAAGATCGGTTATCGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 651
Db 241 GAAGATCGGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
Qy 652 TATGTTATGATGACCAAGATTTATGTTTATGATGATGATGATGATGATGATGATGATGAT 711
Db 301 TACCTTATGATGACCAAGATTTATGTTTATGATGATGATGATGATGATGATGATGATGAT 360
Qy 712 AAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
Db 361 AAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 772 TCAATGTTTATGTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 831
Db 421 TCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 832 GTTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
Db 481 GTTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 892 ACTGTGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
Db 541 ACAGTTGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 952 CAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
Db 601 CAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 1012 ACATTTATCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
Db 661 ACATTTATCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 1072 ACTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
Db 721 ACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

Db 1589796 CAATCATCTGCTCAGTGGAGAGCTGATATACGTATACCTCAATCTTGTCTTA 1589737
 Qy 966 TGCACAATTTGCTTGTGTCGCTAAGAACGCAAGATACGTATATTCCTTA 1025
 Db 1589736 TCGGCCAATTTCTTGTGTCGAGAGCAAGACGTTTGCACAGTTTACAGATAT 1589677
 Qy 1026 CATTAAATGAGCAAGTGAAGAACCATTAAGAACGTTTCAAGTTTAAATGAGT 1085
 Db 1589676 GATTAAAGCTATTACGGAAGCAAGCAAGATTTCCGTTGCTTATTAATGATGT 1589617
 Qy 1086 CGTTATGATGCAATTTCCATGACCCGTAACCAAGCAAGCAAGTGAAGTGT 1145
 Db 1589616 GATTATGATGCTATGCTATTAATCAACCAAGTGAAGCAAGCAAGTGAAGT 1589557
 Qy 1146 CTTTATGCAACACAGTTCGATGATGACCAAGCAAGTTCGTTGTTTAAATGATGT 1205
 Db 1589556 TTTTACGCAACTGCAAGTTCGTTTAAACCAAGTTCGTTTAAACGGA 1589497
 Qy 1206 AGAATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAATCCGTCGCTTTGG 1265
 Db 1589496 AGAATTAATGCACTTTTCAATCTTCTTCTTAAATCAATCCGTAAGCATTTGT 1589437
 Qy 1266 TTTTGAAGTACCAATTCATATTATAGCTCGAAA 1301
 Db 1589436 TTTTGAAGAACCAATTAATCTAATAGACGTAA 1589401

RESULT 14

ABK75008 standard; DNA; 1311 BP.

ABK75008;

13-AUG-2002 (first entry)

Bacillus licheniformis genomic sequence tag (GST) #2299.

Differential gene expression; genomic sequenced tag; GST;

altered culture condition; environmental stress;

physiological provocation; ds.

Bacillus licheniformis.

WO200229113-A2.

11-APR-2002.

05-OCT-2001; 2001WO-US31437.

06-OCT-2000; 2000US-0680598.

27-MAR-2001; 2001US-279526P.

(NOVO) NOVOZYMES BIOTECH INC.

(NOVO) NOVOZYMES AS.

Berka R, Clausen IG;

WPI: 2002-416684/4.

Monitoring differential expression of several genes in first Bacillus

cell relative to expression of same genes in one or more second

Bacillus cells, by using substrate containing Bacillus genomic

sequenced tag array

Claim 4: SEQ ID NO 2299; 200pp; English.

a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterization is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.

Sequence 1311 BP; 376 A; 279 C; 353 G; 303 T; 0 other;

Query Match 45.9%; Score 599.4; DB 24; Length 1311;

Best Local Similarity 66.2%; Pred. No. 5,3e-128;

Matches 864; Conservative 0; Mismatches 441; Indels 0; Gaps 0;

1 ATGACTAAACCTATAGTACTATTTAGTAGGCTTAATGTAGTAAATCTCAATTTT 60

1 ATGGGAAAACCTGTGCTGACCTTACTTGGAGACCCCAATGTGGGGAAGTCCAGATCTT 60

61 AATGAAATAGTTGGAAGACGTTGTTGATTTGTTGAAGATACGCCAGGTGAACAGAGAT 120

61 AACAGATTCGCGGCGGAAAGAAATTCGATTTGGAAGATACACCTGCGTACGCGGAT 120

121 CGTATTTATCTTCAAGTGCAGTGTAAACACATGATTTCAATATATGATACAGTGT 180

121 CGGATATACGCTGCTCGGAGTGCCTGAATCAAGCTTCACTTATGATATGATACAGG 180

181 ATTGAATTTGATGATGACCATTTCCAAACACAAATTAAGAGCCAGGCAAGAAATCCCAT 240

181 ATCGAAGTCGGGATGAGAGCCGTTTGGCGCAGATTTGGCATACAGGCGGAGATCCGCA 240

241 GATGAAGCGGATGTTATTTATTTATGTTAAGTTCGCGTGAAGATTTGACACAAAGCAT 300

241 GAAGAAGCAGATGATATTATTTTCATGACGAAGCGCGGAGAGGCGTCACACCCGCTAT 300

301 GAAATGTCGCTCAATTTTATCAAAATCAAAACCGGCTGATTAAGCGGTTAAACAA 360

301 GAAGAAGTGGCGAAATTTTATACCGTACAAAACCCGCTGCTGGGTTAAATAA 360

361 GTAGATTAATATGAATGCGTACAGAGCGTATGATTTCTATTCAATTAGATTTGGTGA 420

361 GTGATTAATCCCTGAATATGAGACAAACATTTTATGATTTTATGCCCTCGCTTGGCAG 420

421 CCGTATCCAAATATACAGGCTCACATGTTAGTCTTGTGACTTGTAGATGACATGTT 480

421 CCGTCCCGGATTTATGATGACGATGCGCTTGTGCTGCGGATCTGCTGACCGGTCAC 480

481 TCTCATTTTGGGAGAGAGAGAAATCTTATGATGAAGATACAAATTCGATATCCAT 540

481 GAACATTTTAAACAAATATACCCGAGAGAGATGAGATGAAGTCTCAATTTTGGCC 540

541 ATTGAGCAGCAACAGTATGATTAATCAAGTTATGATTAATTTAGTGAAGATGCG 600

541 ATCGCCCGGCGGACAGCTGGGAAATCTTCCCTTTGAACCCGATGATTTGGAGAAAGCGC 600

601 GTTATGTCAGCAATATTTGGGAGAGAGCGGCGGATGATGACACGAGTTTACATAC 660

601 GTTATGTTTCTAATGTTGACAGGACGACGAGAGCGTATTTGATACAGATATAGTAT 660

661 GATGACAAAGATTTATGTTTATTCATATCTGCTGATGCTGATGCTGATGCTGATGCT 720

661 AATCAAAAGAGATTTGATCTGATGATACAGCGGATGAGAAAGGGGAAATGAT 720

721 GAATCACTGAGAAATATTCAGTATTAAGAGCTTTAAAGCATTTGAACGTTCAATGT 780

Db 721 GAGGCACTGAAATAACAGCGTCCTCCGAGCTTTAAAGCGATCGACGCTTGAGAATC 780
 QY 781 GTTTAGTGTATATGACGACAAACAGCATTATGACAAATATAACGTTTGACAGA 840
 Db 781 GTCTGGTCTCTTGGACGCGGACGAAGAGCATTTATGACAGACAAAGCGATTTGCCGC 840
 QY 841 TATGACATGACAAAGGTAAGAGCATGCTGATGTTGCTGTAATTAATGATGATGCTGAA 900
 Db 841 TATGACATGACAAAGGTAAGAGCATGCTGATGTTGCTGTAATTAATGATGATGCTGAA 900
 QY 901 AAGATAGTAAACAGTAAAGTAATTTGAAGATGAGTACGTAAAGATTCGAATTTTGA 960
 Db 901 AAGATAGTAAACAGTAAAGTAATTTGAAGATGAGTACGTAAAGATTCGAATTTTGA 960
 QY 961 GATTATGACAAATTCCTTTGCTGCTGTAAGAACGACAGATTTACGATTTATTC 1020
 Db 961 GATTATGACAAATTCCTTTGCTGCTGTAAGAACGACAGATTTACGATTTATTC 1020
 QY 1021 CCTTACATTAATGACCAAGTGAACCATTAATAAAGCTGTTCAAGTTCACTTTAAAT 1080
 Db 1021 CCTGCGGTCTATGACGCGGACGAACCATTTGATGAGGCTGACACCAATTTCTTAAT 1080
 QY 1081 GAAGTGTCTATGACATTTTCCATGACACCTTACACCAACAGCAAGATGACGTTTG 1140
 Db 1081 GATATCATTTATGATGCGGTGCTATGACACCGGACGCGATTAATGAAACCGGTTA 1140
 QY 1141 AATGCTTTTATGACAAAGTGTGATGACACCAAGCAGATTTGTTGATTTGTTAAT 1200
 Db 1141 AATGCTTTTATGACAAAGTGTGATGACACCAAGCAGATTTGTTGATTTGTTAAT 1200
 QY 1201 GATGATGATTAATGATTTCTTTATTAAGCCTATTAGAGATCAAAATCCGCCGT 1260
 Db 1201 GATGATGATTAATGATTTCTTTATTAAGCCTATTAGAGATCAAAATCCGCCGT 1260
 QY 1261 TTTGATTTGAGGTACCAATTCATTTATGATGCTGCAAGAGA 1305
 Db 1261 TTTGATTTGAGGTACCAATTCATTTATGATGCTGCAAGAGA 1305

RESULT 15

ABA90521

ID ABA90521 standard; DNA: 2365589 BP.

XX ABA90521;

XX 16-MAY-2002 (first entry)

XX Genomic sequence of *Lactococcus lactis* IL1403.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.

XX *Lactococcus lactis* IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or *Lactococcus*

XX *lactis* and related species -

XX Claim 1; SEQ ID 1; 2504pp; French.

CC The present invention is related to a *Lactococcus lactis* nucleotide
 CC sequence (ABA90521) and related proteins (AB53300-AB535621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify *Lactococcus lactis* or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO2001/7334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at http://wipo.int/pub/published_pcl_sequences.
 XX

Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;

Query Match 44.6%; Score 582.4; DB 24; Length 2365589;

Best Local Similarity 66.5%; Pred. No. 3e-123;

Matches 867; Conservative 0; Mismatches 431; Indels 6; Gaps 2;

1 ATGACTTAACCTTATGATGCTTATGATGAGGCTTAATGTACGTAATTTTACAAATTTT 60

Db 761403 ATGAGCTTACCTACAGTACGCTATGCGCTGCAATGTCGAAATTCAGATATTC 761462

QY 61 AATAGATAGTGGAGAGCGTGTGATGTTGGAAGATACGCGAGGTGAACAGAT 120

Db 761463 AACCGTATTCACAGAGAACCGATCTCATTTGTCGAACACTTCCAGGTGTAATCCTGAC 761522

QY 121 CGTATTTATTTCTTACGCTGAGTGTGTTAACAATGATTTCAATATTTATGATGAGTGT 180

Db 761523 CGTATTCATCCACAGAGAGATGTTAACCGTAAATTCATATTCATGATGAGTGTG 761582

QY 181 ATTGAATTTGATGATGACCATTTCCAAACCAATTTAGACGAGCAGAAATTCGCCATA 240

Db 761583 ATTGAATTTGATGATGACCATTTCCAAACCAATTTAGACGAGCAGAAATTCGCCATA 761642

QY 241 GATGAGCGGATGATTTATTTTATGATGATGATGATGATGATGATGATGATGATGAT 300

Db 761643 ACCGAGCAGATGTTATTCATCGCTGTTGTTGACGAGAAACAGATTTACAGATGCCAT 761702

QY 301 GAAATGCTGCTCAATTTTATTAACAATTAACAATTAACAATTAACAATTAACAATTAACA 360

Db 761703 GAAATGCTGCTCAATTTTATTAACAATTAACAATTAACAATTAACAATTAACAATTAACA 761762

QY 361 GTAGATTAATTTGAAATGCTTACAGAGCTGATGATTTCTATTCATTTAGATTTGGTGA 420

Db 761763 GTAGATTAATTTGAAATGCTTACAGAGCTGATGATTTCTATTCATTTAGATTTGGTGA 761822

QY 421 CCGATTCATTAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

Db 761823 CCGATTCATTAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 761882

QY 481 TCTCATTTTGGTGAAGAGAGAGATCCTTATGATGATGATGATGATGATGATGATGATGAT 540

Db 761883 CAAATCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 761939

QY 541 ATTGAGCAGACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

Db 761940 ATTGAGCAGACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 761999

QY 601 GTTATGCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657

Db 762000 GTTATGCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762059

QY 658 TATGATGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717

Db 762060 TATGATGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762119

QY 718 TATGATGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777

Db 762120 TATGATGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762179

QY 778 GTTATGCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837

Db 762180 ATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762239

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:59 ; Search time 3513.16 Seconds
(without alignments)
9028.139 Million cell updates/sec

Title: US-09-815-242-4228

Perfect score: 1305
Sequence: 1 atgaactaacctagtagc.....atatatagctcgaagaaga 1305

Scoring table: IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: em_estbba:*
2: em_estbhum:*
3: em_estlin:*
4: em_estlinu:*
5: em_estloy:*
6: em_estplo:*
7: em_estplo:*
8: em_estplo:*
9: gb_estli:*
10: gb_estli:*
11: gb_estli:*
12: gb_estli:*
13: gb_estli:*
14: gb_estli:*
15: em_estfun:*
16: em_estfun:*
17: em_gss_hum:*
18: em_gss_hum:*
19: em_gss_hum:*
20: em_gss_vit:*
21: em_gss_vit:*
22: em_gss_vit:*
23: em_gss_vit:*
24: em_gss_vit:*
25: em_gss_vit:*
26: em_gss_vit:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	ID	Description
1	486.6	37.3	2662	28	BH770770	BH770770 LMGTAG51
2	143.6	11.0	552	9	A1390536	A1390536 mu83b02.y
3	128.4	9.8	721	12	B0573090	B0573090 B0573090
4	124	9.5	1233	11	AV110834	AV110834 Zee mays

5	118.6	9.1	583	14	CA782602	CA782602 sat30n09.
6	108.4	8.3	431	9	AA215138	AA215138 mu83b02.r
7	99.8	7.6	579	13	B0514303	B0514303 EST621718
8	99.2	7.6	532	13	B0281519	B0281519 WHE3022.D
9	89.4	6.9	395	9	A165410	A165410 mu83b02.x
10	88.4	6.8	745	14	CD377507	CD377507 PTM02676
11	84.6	6.5	171	28	BH234869	BH234869 PE_04.x
12	83.2	6.4	1101	29	CNS00396	AL0630321 Drosoph11
13	81.4	6.2	167	28	BH234795	BH234795 MEA_B03.
14	79.6	6.1	653	13	B0855888	B0855888 OCB28A05.
15	78.8	6.0	501	14	N37967	N37967 10194 Lambd
16	78.8	6.0	613	10	BG299568	BG299568 HVSME002
17	77.4	5.9	554	14	CA938405	CA938405 sav32b06.
18	77	5.9	631	28	BH861059	BH861059 M3B1T7 we
19	76.8	5.9	727	14	CD407144	CD407144 GM_CK3216
20	74.2	5.7	908	28	A2548467	A2548467 ENTEK30TR
21	73.2	5.6	1359	28	BH770563	BH770563 LMGTAG32
22	73	5.6	769	13	B0802107	B0802107 WHE2822.B
23	72.6	5.6	1010	14	CD102897	CD102897 AGENCOURT
24	71.6	5.5	410	14	CD474342	CD474342 nad03-2ms
25	71.6	5.5	841	14	CB996761	CB996761 AGENCOURT
26	71.6	5.5	905	28	A2550256	A2550256 ENTEV58TR
27	71.4	5.5	487	12	B1272296	B1272296 NF022D05F
28	71	5.4	538	13	BQ490035	BQ490035 60-E9427-
29	70.6	5.4	552	14	CA742475	CA742475 w11s-PK0
30	70.6	5.4	623	12	BJ463635	BJ463635 BJA63635
31	69.8	5.3	843	28	A2551618	A2551618 ENTDV54TR
32	69.8	5.3	931	28	BH160272	BH160272 ENTQV49TR
33	69.2	5.3	412	12	BH402985	BH402985 SLA010D07
34	69	5.3	772	12	B1920601	B1920601 EST540536
35	68.8	5.2	849	28	A2546009	A2546009 ENTFW53TF
36	68.4	5.2	912	28	A2551092	A2551092 ENTFJ22TF
37	67.2	5.1	745	28	AY079759	AY079759 AY079759
38	66	5.1	911	28	BH164417	BH164417 ENTS163TR
39	65.8	5.0	532	10	AM979775	AM979775 EST341400
40	65.6	5.0	541	13	BQ627620	BQ627620 SAP33h03.
41	65.6	5.0	877	28	A2531291	A2531291 ENTBO34TR
42	65.2	5.0	571	13	B0578174	B0578174 sar48q07.
43	64.2	4.9	721	10	BG646649	BG646649 EST508268
44	64	4.9	350	14	CA735279	CA735279 wpl1s-PK0
45	63.6	4.9	356	14	H37009	H37009 15138 Lambd

ALIGNMENTS

RESULT 1
BH770770
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BH770770 2662 bp DNA linear GSS 01-MAY-2002
LMGTAG515 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, genomic survey sequence.
BH770770.1 GI:20373727
GSS.
Lactococcus lactis subsp. cremoris
Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
1 (bases 1 to 2662)
Bojoltin,A., Ehrlich,S.D. and Sorokin,A.
Studies of genomes of dairy bacteria Lactococcus lactis
Sci. Aliments. (2002) In press
Contact: Sorokin A
Genetique Microbiene
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is yphL (98%)
Class: Shotgun
High quality sequence start: 30
High quality sequence stop: 2634.

```

FEATURES
  source
    Location/Qualifiers
      1.2662
      /organism="Lactococcus lactis subsp. cremoris"
      /mol_type="genomic DNA"
      /strain="MG1363"
      /db_xref="taxon:1359"
      /clone_lib="MG1363 Random Sequence Tag Library"
      /note="Vector: pSGM2; Site_1: SmaI; Library of
      chromosomal fragments of L.lactis strain MG1363 was
      prepared by partial AluI digestion or by sonication."
BASE COUNT      871 a      506 c      513 g      772 t
ORIGIN
Query Match      37.3%; Score 486.6; DB 28; Length 2662;
Best Local Similarity 65.6%; Pred. No. 5.2e-89;
Matches 743; Conservative 0; Mismatches 384; Indels 6; Gaps 2;

QY      1 ATGACCTAACCTATATAGCTATGTAGAGCGCTTAATGTAGCTAACTACAAATTTT 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1533 ATGAGCTTACCTACAGATCGCATCGTTGGCGCTCAAAATGTTGGAATTCACGATATTC 1592

QY      61 AATGAGATGTTGGAGACGTTGTTGATTTGTAAGATACGCGAGGTGAACAGAGAT 120
      || || || || || || || || || || || || || || || || || || || || ||
Db      1593 AACCGATTATGACGAGAGAACGATCTCAATTTGTGAAGACATTTCCAGGTGTAACCTGAC 1652

QY      121 CGTATTATTTCTTCAGGTAGTGGTTAACACATGATTTCAATATTTATGATACAGTGT 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1653 CGTATCTATGCCACAGAGATGTTAACCCGTAATTCACATCATCGACCGGTGGG 1712

QY      181 ATTGAATGTTGATGATGACCATGCCAACACAACTAATAGACGCGAGCAGAAATCCGCATA 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1713 ATTGAATCTTCATACATACCTTTTCATGATGGAATTCGGCAGCCGATG 1772

QY      241 GATGAAGCGAGTGTATTTATTTTATGTTTATGCTGCGTGAAGATTGACACAAACGAT 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1773 ACCGAGAGATGTTATTCGCTGTCGTTGATGAGAAACAGAGATTACTGATGCCAT 1832

QY      301 GAAATGTCGCTCAAAATTTTATACAAATCAATAAAACCGGTCGTATTAGCGGTTAAACAA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1833 GAAGCAGTTGCTAATATTTCTTTATCTGACAGATAAACCTGTCTTATCGTCAATAAA 1892

QY      361 GTGATTAATATGGAATGGGTACAGACGTGATATTTCTATTCATTAATGTTGGAA 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1893 GTGCAACACCCGAGACGTGATGGAATTTTGTACTTCTATTCAGTGGACGTGGGAT 1952

QY      421 CCGTATCCATATACAGGTCACATGTTAGTCTTGTGTCATCTGTTAGTACAGTGTGT 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1953 CCATATCTGTATCAGCTGTTCACGGGATTGGACAGAGGATGATCTGATGCTATTGTT 2012

QY      481 TCTCATTTTGGTCAAGAGAGAGATCCTTATGATGAGATACAAATTCGACTATCCATT 540
      || || || || || || || || || || || || || || || || || || || || ||
Db      2013 CAAAACTCTCTAATCAATCAAGA---AGAAAAAGATGATCAAGTTCAGCTTG 2069

QY      541 ATTGAGCAGCAACGATGATTAATGATTTAGTAAATGCTATTTTATGTAAGATCCG 600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2070 ATTGCTGCTCAAAAGCTTGTTAATATCTCTTATCAATGCAATCTTGGTGAAGACCGA 2129

QY      601 GTTATCGTTTCTAATGTTGAGGACACAGACGCTATTTGATACAGAGTA---TAGT 657
      || || || || || || || || || || || || || || || || || || || || ||
Db      2130 GTGATTGCAAGCCCAATCGCGGGAACAACGATGATGACCATTTGATACACCTTCGTTGAT 2189

QY      658 TATGATGACACATTAATGTTTAAATGATGCTGTGTTAGCGTTAAAAAGAGAAAGTA 717
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2190 TCTGAGACCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2249

QY      718 TATGATCAACATGAGAAATTTCAATTAATGAGCTTTAAACGATGACGTTCAAT 777
      || || || || || || || || || || || || || || || || || || || || ||
Db      2250 TACGAAACAGTAAATAATTTCAATGATGCGTGCATGCGTCCATGACCGTTGAT 2309

QY      778 GTTGTATTAGTGTATTGATGACAGAAACAGCATTTATTAACAGATTAACGTTGTGA 837
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2310 ATCGTTTGTATGCTCATTAATGCTGAAGAAGATTCGCAATATGATATGCTATCGCT 2369

```

```

QY      838 GGATATGACATGACAGACAGCTAAAGACGCTGATTTGCTGTAATTAATGGCTACTG 897
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2370 GGATTTGCTATGACAGCTAAAGGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2429

QY      898 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2430 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2489

QY      958 TTGATTTATGACAAATTTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2490 CTGATTTATGACCAATTCGTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2549

QY      1018 TTCCCTTACATTAATGACAGACGTAAGACATTAAGACGTTTCAAGTTAACTTA 1077
      || || || || || || || || || || || || || || || || || || || || ||
Db      2550 CCAGATATGATGATTAAGAAATTCACATGCTCAACAAACCTGATTTCAAGTTGTTG 2609

QY      1078 AATGAGTGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2610 AATGATGATGATTAATGATGCTGTTGCGATTAATCACTCCACAGATTAAGG 2662

```

```

RESULT 2
AI390536
LOCUS      552 bp      mRNA      linear      EST 15-MAR-2000
DEFINITION      mu83b02.y1 Stragene mouse melanoma (#937312) Mus musculus CDNA
clone IMAGE:652107 5' similar to SW:YPHC_BACSU P50743 HYPOTHETICAL
48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION. ; mRNA
sequence.
ACCESSION      AI390536
VERSION      AI390536.1 GI:4216543
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      Eukaryota: Chordata: Craniata: Vertebrata; Euteleostomi;
Mammalia: Eutheria: Rodentia; Scurionathu; Muridae; Murinae; Mus.
1 (bases 1 to 552)
AUTHORS      Maria M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, T., Person
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Riller
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterson, R., and Wilson, R.
The WashU-NCI Mouse EST Project 1999

```

```

TITLE      Unpublished
JOURNAL      Contact: Maria M/WashU-NCI Mouse EST Project 1999
COMMENT      Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 420
POLYA-No.

```

```

FEATURES
  source
    Location/Qualifiers
      1.552
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /clone="IMAGE:652107"
      /tissue_type="melanoma"
      /dev_stage="M2 cells"
      /lab_host="SOLR (kanamycin resistant)"
      /clone_lib="Stragene mouse melanoma (#937312)"
      /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
      ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
      dT. From M2 cells, a highly metastatic derivative of the
      K-1735 (mouse) melanoma. Average insert size: 1.0 kb.

```

BASE COUNT	220 a	63 c	94 g	174 t	1 others
ORIGIN	Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCCACGAG				
	3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'				

Query Match	11.0%	Score 143.6	DB 9	Length 552
Best Local Similarity	56.18	Pred. No. 4.0e-19		
Matches 316; Conservative	0	Mismatches 235	Indels 12	Gaps 2

Oy 136 GGGAGTGTCTTAACATGATTCATATTTATTCATACAGTGTGATTTGAATTCGTGAT 195
 | | | | |
 Db 1 GTTGAATTAGACAGAAAGGAATTTTATTCATCATGACACAGCGGAATTTGAAATGAAAC 60
 | | | | |

Oy 196 GGCACATTCCTCAACACACAAATTACAGCGCAGCAAAATCCGCATAGATGAAGCGGATGTT 255
 Db 61 AATCATTTTCAAGAACCAATTAGAAATTTCAACTTCACAAATGCAANTGAAAGCAAAATTTTA 120

Oy 256 ATTATTTTATTGGTTAACGTGCGTGAGAGATTACACACAAGCATGAAATGCTGCCTCAA 315
| | | | | | | | | | | | | | | | | |
Db 121 ATTTATTTTTTAGTAGATCGTAAGTTGAATAATGATTTCAGATGATCATTGTTGTTATAGAT 180

Qy 316 ATTTTATCAAACTAAAAAACCCTGCTATACCGGTTAACACAAGTAGATATATGCA 375
Db 181 TTGCTTCGAAAAATCAAGCAAAAAAGTTTAAATCGCAGCATATTAACATGAAGCAACAA 240

OY	376	ATCCGCTACAGACGCTGATGATTTCTATTCATTTGGATTTGGTGAACCCGATCCAAATATCA	435
Db	241	TTTTTT-----GATACATTCAATTTATTTACACTAGGTTGGTGAACATATATTTTCCAAATTTCA	294

OY 436 GGGTCACATGCTTTAGGTCCTGGTGCACCTGTGTAGATGCAGTTGTTTCATTTTGCTGA 495
| | | | |
Db 285 GCTATTCTATGGACAGACTGTGCGATTATTTATGATGAACAACATTAGAAGTTATGATTTTT 354

QY	496	GAGGAGAAGATCCTTATGATGAGAATCAATTGCACATCCATTATGGACGCCAAC	555
Db	335	ACAAGA-----GAAGAAAATGAAGCATTTAGGTGGCAATTATAGAAAAACC AAC	408

09 556 GTAGTAAATCAGCTTTAGTAAATGCTATTTTGGTGAAGATCGCGTTATCGTTCTAT 615
 409 GCTGGAAATCTACTTTATTAAATAAATTAGCAATGAANAATCGCTCCATTGTTCTCCA 468

0y 61.6 GTTGCAGGCGACACAGAGAGACGGTATTGTATACAGAGTAGTTATGATGTGACCAAGATTAT 675
| | | | | | | | | | | | | | | | | |
Db 469 ATAGCAGGAACAACACGCTGATTAGTTCTAGTTTATTAATAATTTGACCAAAATAGATTNT 528

Qy	676	GTTTAATCGATACGCTGGTAT	698
Db	529	GAATTTATAGATACGCTGGAAT	551

RESULT 3
B573090/c

DEFINITION
BU573090 Ipomoea nil mixture of flower and flower bud
ACCESSION
BU573090 jm22ml3 3', mRNA sequence.

KEYWORDS	EST.
SOURCE	
ORGANISM	
	Ipomoea nil (Japanese morning glory)
	Ipomoea nil

REFERENCE
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Convolvulaceae; Ipomoea.
1 (Phases 1 to 721)

TITLE
JOURNAL
COMMENT
Nitsaka, E., Shinozaki, K., Hayashizaki, Y., Kohara, Y. and Iida, S.
Ests of Japanese morning glory
Unpublished
Otsu, Japan, Chitose
Otsu, Japan, Chitose

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

FEATURES
Source Fax: 81-559-81-6855
Email: tsuhin@genes.nig.ac.jp.
Location/Qualifiers
1. .721

```

/mol_type="mRNA"
/cultivar="Tokyo-kokel standard"
/db_xref="taxon:35883"

```

BASE COUNT		
186 a	/tissue_type="mixture of flower and flower bud"	
184 c	/clone_11b="Ipomoea nil mixture of flower and flower bud"	
		136 g 215 t

Query Match	9.88;	Score 128.4;	DB 12;	Length 721;
Best Local Similarity	52.88;	Pred. No. 6.2e-16;		

587 TAGCTGAAGATCCGCGTTATCGTTTCTTAATGTTGCAGGCACACGAGACGCTATTGATA 646

647 CACAGTATA---CTTATGATGGACAGATTATGTTTAAATCGATCTGCTGTAAGCGTA 703

QY 704 AAAAAGCAAGATATATCATC-----AACTGACAATAATTACAGTATTAGAGCTT 754
||||| |||| | |||| | |||| | |||| |

QY 755 TAAAGCGATGCAAGCTTCATGCTGTTTACGCTATTGTCAGACAGACAAGCATTA 814
+ ||| ||||| ||||| ||| ||||| || |||||

QY 815 TTGCACACAGATAAACGTC^{*}TTGCACAGATATGCACATGAACACAGCTAAAGCAGTCGTGATTGG 874

875 TCGTAATTAATGGGATCTGTGCAAAAAGTACTAA--AACGATGACAAATTTGCAACA 932

933 TGAGTACGTAAAGCATTTCCCAATTTTACATATGACACAATTCGTTTGTGTCGTAA 992

993 AGAAGCCACAGATTACGTACATATTCCCTTACATATATGAAGCAAGTGAACCATTA 10522

1053 AAAACGTGTCAAAGTTCACATTTAAATGACGCTACTGATGCATTTCCATGACCC 1112

QY 1113 TACACCAACAGACAAAGT--AGACGTTGATGTCCTTTATGACACACAAAGTTGCTAT 1169

QY 1170 AGAACCAACGACATTGTGTATTTGTTAATGATGATAGATTAATGCATTTTCTATAA 1239

QY 1230 ACAGCTATTTAGAGATCAATCCCGCTTTTGCTTTGAGAGTACACCAATTC 1285

RESULT 4

LOCUS	AY110834	1233 bp	mRNA	linear	HTC 17-OCT-2002
DEFINITION	Zea mays CL915_1 mRNA sequence.				
ACCESSION	AY110834				
VERSION	AY110834.1	GI:21215424			

KEYWORDS HNC.
SOURCE zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Halney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1233)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZMDB and may be found by BLAST
searching at MSU, maizegap.org; ZMDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schubler, Iowa State, then clones may be requested from ZMDB:
www.zmdb.iastate.edu.

FEATURES
SOURCE Location/Qualifiers
1..1233
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:633068"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/Dupont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont configs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

BASE COUNT 335 a 268 c 294 g 278 t 58 others

ORIGIN
Query Match 9.5%; Score 124; DB 11; Length 1233;
Best Local Similarity 51.1%; Pred. No. 5.2e-15;
Matches 362; Conservative 0; Mismatches 328; Indels 18; Gaps 4;

QY 616 GTTGCGAGCAACGAGAGCGCTATTGATACAGATAGT--TATGATGCAAGAT 672
||| ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| |||
Db 21 GTTAGTGGGACACAGCTGATGCCATGTACTGATGAGTACTACAGAGATGGGAGAAG 80
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 673 TAGTTTAAATCGATCTGCTGCTGATCCGTAAGAAAGAAAGTATAT-----GAA 723
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 TACAACTCATTTGATCTGCTGGGATCCGGCGGAGAGACAGCTTATTTCTGCGCAGC 140
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 724 TCAAGTGAATATTTAGATATTAGAGCTTTAAAGCGATTTGAAGCTTCAATGTTGT 783
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 ACAACCGAATCATCTTGGTAAAGCGTGCAATTCGACAAATGGCCCTCGATGCTGTT 200
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 784 TTAGTGTATTGATGAGCAAGAGCATTTATGAACAAGATAAAGCTTTGACAGATAT 843
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 GCCCTGTATTGAGCAATGGCCGTGTGCACAGAGAGATTTAAATTTGNNNNNNN 260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 844 GCACATACACAGGTAAAGAGTCGTGATTTGCTAAATAAATGGATCTGTGAAAAA 903
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 NNNNNNNNNNNNNNNNNNNNTTGTGTCATTTGTTGAACAATGGATACATCCCAAC 320
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 904 GATAGTAAACGATGAGA---AATTGAGATGATAGCTAGTAAAGAAATCCATTTTGA 960
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 AAGAACCGAGAGATACACACATTTATGAACAAGATTAAGAGAGCTTCGCATCTT 380
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 961 GATTATGCAAAATTTGTTGTGCTGCTAAAGACGACAGATTAAGTACATTTATTC 1020
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 GATTGGGACCTATTTGCTACTGTTCTGCGAACAATGGCACCACTGTTGAAGAATAT 440
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1021 CCTACATTAATGACAGAGTGAAGAAACCATTAAGAAACGCTGTCAAATTCATTTAAT 1080
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 TCTGCTCTGCTCTTTTGGTGAAGAAAGGTCTAGAGACTCAGACCTCATTTCTTAAT 500
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1081 GAATGCTTTGCTAGTGAATTTTCATGACCCCTACACCAAGCAAGT---AGACGT 1137
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 CAACTTTATTAGAGAGCTATTAAGCTTCAATTCACCAACCAAGCAAGTGGCAAGA 560
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1138 TTGAATGCTTTTATGACACAGAGTTCATATAGAACCCGACATTTGTTATTTGTT 1197
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 GATGCTGTATTTATACACACAGAGCTGCGTGCCTGACCAACATTCGTTCTATTGTA 620
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1198 AATGATGATTAATTAATGATTTTCTTATTAAGCTATTTAGAGATCAATCCGCGC 1257
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 AATGATGCAAGCTCTTCCCTGACACCTACCGCGTTATGCAATGAAAGAACTCGGTCA 680
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1258 GCTTTGCTTTTGAAGGTACACCAATTCATTTATTTACCTGGAAGAGA 1305
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 681 GACCGCGTTTCCAGCACCCTATTTCGTTACTGTGCGGTAGCAGA 728
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
CA782602
LOCUS
DEFINITION sat30b09.y1 Gm-c1056 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1056-3065 5' similar to TR:P2548 P72348 HYPOTHETICAL PROTEIN
;', mRNA sequence.
ACCESSION CA782602
VERSION CA782602.1 GI:26044687
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Epeiding, J., Coryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., Mccann
R., Waterston, R. and Wilson, R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 421.
Location/Qualifiers

FEATURES
SOURCE

1..583
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1056-3065"
/tissue_type="Whole seedling, 4 day old"
/lab_host="DH108"
/clone_lib="Gm-c1056"
/note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 4 day old seedling of P1468916. The seedlings were
germinated in a growth chamber using germination paper.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly(dT) sequence with a XhoI restriction

[illegible]

QY	879	AAATAAAGGATCTACTGGGAAAAAGATGATGAACGATTGAAGAATTGTAAGTGAAGT	938
Db	381	GACACAAGTGCGATTACGATCCCAACAGAACCAAGAACTACTGTATTTTTAGAGAGA	440
QY	939	ACGTAAA	945
Db	441	TGTTAGA	447
RESULT 8			
BQ281519			
LOCUS		532 bp	mRNA linear EST 13-MAY-2002
DEFINITION	WHE3022_D01.G02Zs wheat unstressed seedling shoot normalized cDNA library Triticum aestivum CDNA clone WHE3022_D01.G02, mRNA		
ACCESSION	BQ281519		
VERSION	BQ281519.1		
KEYWORDS	GI:20549073		
SOURCE	EST.		
ORGANISM	Triticum aestivum (bread wheat)		
REFERENCE	Triticum aestivum		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum.		
JOURNAL	Anderson O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T., Rausch,C.U., Seaton,C.B., Tong,J.C. and Zhang,D.		
COMMENT	The structure and function of the expressed portion of the wheat genomes - Normalized shoot cDNA library Unpublished Contact: Olin Anderson US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center 800 Buchanan Street, Albany, CA 94710, USA Tel.: 5105595773 Fax: 5105595818 Email: oanderson@pw.usda.gov Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: SK primer. Location/Qualifiers 1..532 /organism="Triticum aestivum" /mol_type="mRNA" /cultivar="Chinese Spring" /db_xref="taxon:4565" /clone="WHE3022_D01_G02" /tissue_type="Etiolated shoot" /dev_stage="Five day old seedling" /lab_host="E. coli DH10B" /clone_lib="wheat unstressed seedling shoot normalized cDNA library" /note="Vector: Lambda Uni-ZAP XR, excised phagemid pluescript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the TJ Clase lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give pluescript phagmids before normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares'. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."		
BASE COUNT	162 a	112 c	110 g 148 t
ORIGIN			
Query Match	7.6%	Score 99.2;	DB 13; Length 532;

Best Local Similarity 53.3%; Pred. No. 5.5e-10;
Matches 256; Conservative 0; Mismatches 218; Indels 6; Gaps 2;

QY 753 TTTAAAGCGATGACCTGCAATGTTGTTTATTGCTGTTATGTCAGAAAGGCAAT 812
 DB 48 TTTTCTCAATTCGTCCTGCTGATGTTGTTCCCTGTTATCTAGCAAGGCGCTGTCT 107
 QY 813 TATTAAGCAAGTAAAGCTGTTGAGATATGTCATGAAAGGTAAGGAGTGTGAT 872
 DB 108 TACGACAGCAAGATTTTAAATTTGTCACAGAGATTCAGCAAGAAAGGCTTGTCTAT 167
 QY 873 TGTCTAAATTAATGATGATCTGTTGAAAAAGATA--GTAAAGATGAAAGAAATTTGA 929
 DB 168 TGTGTGAACAATGGATGATCTATCCCTAACAGAACATGAGATACAAACATATATGA 227
 QY 930 AGATGAAGTACGTAAGATTCGCAATTTTATGATATGACAAATTCCTTTGTCTGTC 989
 DB 228 ACMAAGATGATATAGAAAGCTTCCTCATCTGATGAGGACCTATGCTGTTCTGTC 287
 QY 990 TAAAGACGACAGATTAAGTACGATATTCCTTACATTATTAAGCAAGTGAAGCA 1049
 DB 288 GACTATATGACACACGATGTCACAAAAGATATTTCTGCTGCTGTTGTTGAGAAATGAAG 347
 QY 1050 TAAAGACGCTGTTCAAGTTCACCTTTAATGAGTGTCTGATGCAATTTGCATGAA 1109
 DB 348 GTCTATTAAGACTGCGACCTCCATCTTATCATGATATTAAGAAAGTATGCAATTCGA 407
 QY 1110 CCCTACACCAAGCAAGAAAGT--AGACGTTGATGCTCTTTATGACACCAAGTTCG 1166
 DB 408 ACCACCAACCAAGCAAGAGGTCGCAAAAAGAGTCGCTATTTATACAAACAGAGCTGC 467
 QY 1167 TATAGAACCCGACGATTTGTTGTTTATGATGATGATGATTAATGATTTTCTTA 1226
 DB 468 CATGTGTCACCAACATTCGTTTATTTATTTGTAATGATGAGAAAGCTTCCTGATACATA 527

RESULT 9
 A1465410/c 395 bp mRNA linear EST 09-MAR-1999
 LOCUS mu83b02.x1 Striatogene mouse melanoma (#937312) Mus musculus CDNA
 DEFINITION clone IMAGE:652107 3' similar to SW:TPHC_BACSU P50743 HYPOHETICAL
 48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION. // mRNA
 sequence.

ACCESSION A1465410
 VERSION A1465410
 KEYWORDS EST
 SOURCE A1465410.1 GI:4319440
 ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 395)

AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person,
 B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rilter,
 E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.

TITLE The WashU-NCI Mouse EST Project 1999

COMMENT Unpublished
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

FEATURES
 source 1..395
 location/Qualifiers
 High quality sequence stop: 276.
 Possible reversed clone: similarity on wrong strand

1..395
 location/Qualifiers
 High quality sequence stop: 276.
 Possible reversed clone: similarity on wrong strand

/mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:652107"
 /tissue_type="melanoma"
 /dev_stage="M2 cells"
 /lab_host="SOLR (Xenopus resistant)"
 /note="Jb-Stratogene mouse melanoma (#937312)"
 /clone_organ="Stratogene mouse melanoma (a highly metastatic derivative of the
 K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
 dr. From M2 cells, a highly metastatic derivative of the
 K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
 3' -3' adaptor sequence: 5' GATATCGGACGAG
 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'")
 BASE COUNT 140 a 41 c 47 g 167 t

Query Match 6.9%; Score 89.4; DB 9; Length 395;
 Best Local Similarity 55.3%; Pred. No. 5.3e-08;
 Matches 194; Conservative 0; Mismatches 156; Indels 1; Gaps 1;

QY 955 TTTTATGATTAAGCAACATTCGTTTGTGCTGCTAAAGAACGCAAGATTAAGTACA 1014
 DB 364 TTTTATGAGCTGAGCCCAATTTGTTTATTTTCAAGC-ACAACAGGAAGTGAATTTCTTAA 306
 QY 1015 TTTATCCCTTACATTAATTAAGCAAGTGAAGCAACATTAAGAAAGCTGTCAAGTTCACT 1074
 DB 305 TTAAGAGAACCAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
 QY 1075 TTAATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
 DB 245 CTTAATCAATTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
 QY 1135 CGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 DB 185 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 126
 QY 1195 GTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
 DB 125 GTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
 QY 1255 GCCGCTTTGTTTGAAGTACACCAATTCATTAATTAATTAATTAATTAATTAATTAATTA 1305
 DB 65 AACATATTTGTTTTCAGCTGTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15

RESULT 10
 CD377507 745 bp mRNA linear EST 31-MAY-2003
 LOCUS PTTM02676 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cdna 5', mRNA sequence.
 ACCESSION CD377507.1 GI:31253121
 VERSION CD377507.1
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 745)

AUTHORS Scala, S., Carels, N., Falciatore, A., Chiusano, M. L. and Bowler, C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 22111123
 PUBMED 12114555

COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: christel@phn.szn.it
 Diatom EST Database (http://avesstagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA-No.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 17:17:09 : Search time 5327.98 Seconds
(without alignments)
10020.134 Million cell updates/sec

Title: US-09-815-242-4228

Perfect score: 1305

Sequence: 1 atgactaacctactagtc.....atatctagctcgaagaga 1305

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_pl:*

8: gb_pr:*

9: gb_ro:*

10: gb_sts:*

11: gb_sy:*

12: gb_un:*

13: gb_vl:*

14: em_ba:*

15: em_fun:*

16: em_hum:*

17: em_in:*

18: em_mu:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_pl:*

24: em_ro:*

25: em_sts:*

26: em_un:*

27: em_vl:*

28: em_hcg:*

29: em_hcg_inv:*

30: em_hcg_other:*

31: em_hcg_mus:*

32: em_hcg_pln:*

33: em_hcg_rnd:*

34: em_hcg_mam:*

35: em_hcg_vrt:*

36: em_sy:*

37: em_hgtg_hum:*

38: em_hgtg_mus:*

39: em_hgtg_other:*

40: em_hgtg_pln:*

41: em_hgtg_rnd:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1292.2	99.0	301550	1	AP003134 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
2	1292.2	98.9	346900	1	AP003362 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
3	1290.6	98.9	1308	6	AX62268 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
4	1290.6	98.9	333750	1	AP004827 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
5	919.4	70.5	3269	1	AF270032 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
6	919.4	70.5	3269	6	AX145350 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
7	919.4	70.5	300892	1	AE016747 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
8	733.8	56.2	304680	1	AE017002 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
9	727.4	55.7	290117	1	AE017028 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
10	683.2	52.4	347050	1	AL591981 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
11	683.2	52.4	349980	6	AX641670 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
12	676.8	51.9	313450	1	AL596170 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
13	676.8	51.9	319630	6	AX413016 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
14	676.8	51.9	349980	6	AX417046 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
15	667.6	51.2	960	6	AX144037 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
16	656.8	50.3	300050	1	AP004599 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
17	647.4	49.6	24887	1	BACSER1 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
18	647.4	49.6	213680	1	BSUB0012 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
19	634.8	48.6	302173	1	AE016951 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
20	633	48.5	300550	1	AP001512 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
21	600	46.0	1311	6	AX607165 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
22	600	46.0	20601	1	AE014265 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
23	600	46.0	44145	6	AX602195 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
24	600	46.0	174050	1	SAG766852 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
25	599.4	45.9	1311	6	AX433884 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
26	582.4	44.6	11071	1	AE006309 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
27	552	42.3	52276	1	AE014141 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
28	552	42.3	151947	2	SPNEU1902 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
29	552	42.3	323825	1	AP005146 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
30	550.4	42.2	1308	6	AX570314 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
31	550.4	42.2	5066	6	BD003759 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
32	550.4	42.2	10310	1	AE007464 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
33	550.4	42.2	349980	6	AX571764 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
34	550.4	42.2	349980	6	AX571765 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
35	550.4	42.2	349980	6	AX571765 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
36	549.6	42.1	12434	1	AE006498 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
37	548.8	42.1	12370	1	AE009978 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
38	548.8	42.1	12340	1	AE008523 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
39	544.2	41.7	302050	1	AL935257 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
40	531.2	40.7	3737	1	AB016077 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
41	531.2	40.7	13860	1	AE015016 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
42	486.2	37.3	301278	1	AE015939 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
43	476	36.5	296750	1	AP003191 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
44	438.4	33.6	10861	1	AE007680 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
45	423.2	32.4	10691	1	AE010530 Staphylococcus aureus subsp. aureus N315 genomic DNA, complete

ALIGNMENTS

RESULT 1
LOCUS AP003134/c
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS SOURCE
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE 1
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Itoh, J., Ito, T., Kanamori, M.,
Matsunaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,

TITLE Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C., Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J., Kanehisa,M., Yamashita,A., Oshino,K., Furuya,K., Yoshino,C., Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*

JOURNAL Lancet 357 (9264), 1225-1240 (2001)

MEDLINE 21311952

PUBMED 11418146

REFERENCE 2 (bases 1 to 301550)

AUTHORS Director-General, Biotechnology Center, Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.

TITLE Direct Submission

JOURNAL Submitted (30-JAN-2001) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bio@nte.go.jp, URL:http://www.bio.nte.go.jp/, Tel:81-3-3481-1933, Fax:81-3-3481-8424) On Jun 12, 2001 this sequence version replaced gi:13701258.

COMMENT On Jun 12, 2001 this sequence version replaced gi:13701258.

FEATURES

source 1. 301550

organism="Staphylococcus aureus subsp. aureus N315"

/mol_type="genomic DNA"

/strain="N315"

/sub_species="aureus N315"

/db_xref="taxon:158879"

complement(249..947)

/gene="SA1293"

complement(249..947)

/gene="SA1293"

/note="ORFID:SA1293"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BAB4255.1"

/db_xref="gi:13701259"

/translation="MSIFTMITIVYILMLPLMAOHKVKSNYEKYSQVRSCTGTR
EVALEILHANGLYDEVVKGEGFLTDHDPKKKIVSLSPANDRPSVAGTAAAEVG
HAIDHQGVWFLFRALVPAVNLGSSLSYIMIGIILTAIGSAFGSTALMGILM
SLAVLFIIVTLPEFEDASSRAMKQITALINVEKEYKHARKVLSAAMTYAATVAV
AELVILTIARSDD"

complement(1001..1588)

/gene="SA1294"

complement(1001..1588)

/gene="SA1294"

/note="ORFID:SA1294"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BAB4255.1"

/db_xref="gi:13701260"

/translation="MTINTFMOYLYLQSRWMLMLLIGNLGMIVGYVYGEOLSHTPW
HKIFVPDSPATILVTLVLSILIKONKSIIEALAPTLFEKYGIMAVIMNLFTIEO
GDIYNGLVMSHSIMAVOAIVYPRFKRSMIGISVAMIVWFLNDYIDYFIHQPPY
DFTTHWQIGVLSCLSYFGLLYIELNKLKCK"

complement(1578..2153)

/gene="SA1295"

complement(1578..2153)

/gene="SA1295"

/note="ORFID:SA1295"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BAB4255.1"

/db_xref="gi:13701261"

/translation="MSEFLNOIKESFIEYLLFQYRFSRIANVYNIKYNAKLANI
HFVDTKINHHETLEAEVGSASAIQFTKRNLKNTNIEIDYIANHMCAPDIQTHFA
NVSKREORLDDILVAQLTESPSYQYTLHNSMALIDRKHALIDYLLAHNIDSLQNM
EKORFYQLTILNTLKLNVKHNQFEDLADDD"

complement(2167..3411)

/gene="SA1296"

complement(2167..3411)

/gene="SA1296"

/note="ORFID:SA1296"

/codon_start=1

/transl_table=11

/note="ORFID:SA1296"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BAB4255.1"

/db_xref="gi:13701262"

/translation="MEDYIKLIDIDINLQKLEMLDSRVNAITTDNDALFILGETLYN
FGIMPQGLEFVRVLYHKYPDESELLIYFIEGLMSKQDEALEYLSYVEPSEKLMLE
ADLYOINMMEVAIDKLOALELEPNDDPIHFALAMELYDQYLRATSEYETLETG
EYOVNGVNLFSRMADCSLOSNGVSDAIRLYDENDEMTSEDLKKAISYEDNITQO
AIKIMTTLTKRDPDYTOGYLYLOSLYENKRNPDALIEGKEGLRISQFYKELMYTGC
LEIEHGDANESVLLKQALEVNAQOPELLILSDLYRNEDEYEAITTELTYYDEDDO
PTFMWHLAFAITQESERDKAQHFELAPTKNTINDPSDYIFYLIEIGQKQAITTL
NOLLEPSENMWHDLSRLQS"

complement(3418..4716)

/gene="aroA"

complement(3418..4716)

/gene="aroA"

complement(3418..4716)

/note="ORFID:SA1297"

/codon_start=1

/transl_table=11

/product="3-phosphohikimate 1-carboxyvinyltransferase"

/protein_id="BAB4255.1"

/db_xref="gi:13701263"

/translation="WVSEQIIDISGPLAGEIEVPGDKSHTRAIMLASIAGTSNIYK
PLLGEDCRRTMDIFRLIGVDIKEDDKLVNVPQYKAKRTPQVLYTQNSGTTTBLLA
GILSGIDIESVLDGVSIGKRPMDVLRPLKIMANDIEGIDENYPLIIPKPSVGIN
YOMEVAQOVKSAILEASLFSFNDTVYIKELVDSRNHTETMEFHNPIIEAEELSTTT
PDALIOHKPADHPVGDISSAEFFVAALITPESDVTIHNWGINTPRLGIIIDIVKMG
GNIOIFNOTGAEPTASIRIOTYTPMLQPTIEGELVPAKIDELPYALICQAVCTST
IKDAEELKKEPNRIDTTADMLNLGFELOPTINDGLIHPSEFKTNATVYDSLITDRIG
MMLAVASLSSSEPVRIKQPDANVVSFPGFLPKLLENBG"

complement(4726..5790)

/gene="arob"

complement(4726..5790)

/gene="arob"

complement(4726..5790)

/note="ORFID:SA1298"

/codon_start=1

/transl_table=11

/product="3-dehydroquinate synthase"

/protein_id="BAB4255.1"

/db_xref="gi:13701264"

/translation="MKLOTTPSSNNYPIFVEGALDHSITYIDQFDSFILDEVNO
TFADKFDDLSTENHAKVITPAGERTKTFEOYQETLEYILSHHTVNRNALIAGVGAT
GDPAFVAATILRGVHFIOVPTTILAHDSVGKVGINSKQKNDLIGAFRPTAVIYD
LDLEKTLPEEQILSGVYAEVYKHALNGESTQDEIQHKRDKREILGSLGMDKYIAKGI
ETKLDIVVADEKEGCVRRKFLNGLFEGHAEVYNNKISGHAVMIGIITIOFYANILFN
SNHDIQHYINVTYLRGYPLETTTDDIFETIYQYMLSKKNDKQGVQWMLIKHFGDIIV
QHIDQTTLQAHACEQLKTYFK"

complement(5816..6982)

/gene="aroc"

complement(5816..6982)

/gene="aroc"

complement(5816..6982)

/note="ORFID:SA1299"

/codon_start=1

/transl_table=11

/product="chorismate synthase"

/protein_id="BAB4255.1"

/db_xref="gi:13701265"

/translation="MRYLTSGESHGSPOLYIVEGIPANLEIKVEDINKEMPRGCGYG
KGRMQIEKDYIEYISVRNGTTLSPITTVYTNDDFTHKRIKGAAPISSEERNNMK
RTTTPRGHADLVGGMKYNNRDLNRVNERSSARAKRAVAAGALCKVLLDOLDIDY
SVVEIGIKDKDFDSETFKANDLRNDRVYRDSIAQAMRKLIDEANEGSIGGV
GVVVENMPYGVGSVYHVRKLDGKIAQGVGSINARKGVSGFGGFAARPGSEIDGFI
LYNSEIYVGRGSHLGGEGSGNMPITIVNGMKPIPLVYPLNSVDINTKDEKAT
IERSSCAVPASISYCEHYVAFEIKALLEEQSHHIDLOQIDRRQOLNVEF"

complement(7368..7568)

/gene="SA1300"

complement(7368..7568)

/gene="SA1300"

complement(7368..7568)

/note="ORFID:SA1300"

/codon_start=1

/transl_table=11

```
/product="hypothetical protein"
/protein_id="BAB42560.1"
/db_xref="GI:13701266"
/translation="MKNIISIIIGILMFLKMLLGYCAIFLDRPLNPITKIIIFILTL
YIFVLYKELIIFKSKYKNSA"
complement(7597..7722)
/gene="SAS045"
/complement(7597..7722)
/gene="SAS045"
/note="ORFID:SAS045"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAB42561.1"
/db_xref="GI:13701267"
/translation="MMLVIMIIHNKARFLCCSKLLEKFSFLVTRIKKRIFFIK"
complement(7777..8226)
/gene="ndk"
/complement(7777..8226)
/gene="ndk"
/note="ORFID:SA1301"
/codon_start=1
/transl_table=1
/product="nucleoside diphosphate kinase"
/protein_id="BAB42562.1"
/db_xref="GI:13701268"
/translation="MERFPLMKPDVORNLIGFVIRIERKGIKLVGKRMQVPMEL
AEHYGBHCKPEYNDLISFTSAPVPMYVEGEDAVNVRHITGSRPSASFGSIR
GDGLIYGRNIIHOSDSLESAREINLMFENETYSASPRDAWLYE"
complement(8318..9277)
/gene="gerCC"
/complement(8318..9277)
/gene="gerCC"
/note="ORFID:SA1302"
/codon_start=1
/transl_table=1
/product="heptaprenyl diphosphate synthase component II"
/protein_id="BAB42563.1"
/db_xref="GI:13701269"
/translation="MAKLNNNEIKKYEORLEKIRKSDSVLEBASLHLSSGKRVR
PAFVILSSOFKDEQTSQYQVAVALLELHMATLVHDVDYIDKDKRRGKLTISKMD

Query Match      99.0%; Score 1292.2; DB 1; Length 301550;
Best Local Similarity 99.4%; Pred. No. 1.3e-214;
Matches 1297; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGCTATTGTAGTGAGCCCTAATGTAGTAATCTACAATTTT 60
    |||||||
DB 13778 ATGACTAAACCTATAGCTATTGTAGTGAGCCCTAATGTAGTAATCTACAATTTT 13719

QY 61 AATGAAATAGTTGGAGAAGCGTTTCGATTGTTGAAGTACGACGAGGTGAACGAGAT 120
    |||||||
DB 13718 AATGAAATAGTTGGAGAAGCGTTTCGATTGTTGAAGTACGACGAGGTGAACGAGAT 13659

QY 121 CGATTTATCTTCAGGTGAGTGGTTAAACATGATTTCAATTAATTATGATACAGTGGT 180
    |||||||
DB 13658 CGATTTATCTTCAGGTGAGTGGTTAAACATGATTTCAATTAATTATGATACAGTGGT 13599

QY 181 ATTGAATTTGGTATGACCACTTCCAAACACAAATTAAGCGGACGAGAAATGCCATA 240
    |||||||
DB 13598 ATTGAATTTGGTATGACCACTTCCAAACACAAATTAAGCGGACGAGAAATGCCATA 13539

QY 241 GATGAACGCGATGTTATTTTATGTTTAAAGTGGCGTGAAGATTGACACAAAGCAT 300
    |||||||
DB 13538 GATGAACGCGATGTTATTTTATGTTTAAAGTGGCGTGAAGATTGACACAAAGCAT 13479

QY 301 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAACGCGTCTATTAGCGGTTAACAA 360
    |||||||
DB 13478 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAACGCGTCTATTAGCGGTTAACAA 13419

QY 361 GTAGATAATATGAAATGCGTACAGACGTGTATGATTTCTATTATTAAGATTGGTGAA 420
    |||||||
DB 13418 GTAGATAATATGAAATGCGTACAGACGTGTATGATTTCTATTATTAAGATTGGTGAA 13359
```

```
QY 421 CCGTATCCAAATATCAGGTCACATGCTTTAGGTCTTGGTGACTTTAGATGACAGTGT 480
    |||||||
DB 13358 CCGTATCCGATATCAGGTCACATGCTTTAGGTCTTGGTGACTTTAGATGACAGTGT 13299

QY 481 TCTCATTTTGGTGAAGAGAAAGAGATCCTTATGATGAAGATACAAATTCGATTCAT 540
    |||||||
DB 13298 TCTCATTTTGGTGAAGAGAAAGAGATCCTTATGATGAAGATACAAATTCGATTCAT 13239

QY 541 ATTGGACGACCAAAACGTAAGTAAATCAAGTTTAAATGCTATTTTGGTGAAGATGCC 600
    |||||||
DB 13238 ATTGGACGACCAAAACGTAAGTAAATCAAGTTTAAATGCTATTTTGGTGAAGATGCC 13179

QY 601 GTTATCGTTCTTAATGTTGACAGGACCAACGAGAGAGCGTATTGATACAGATATGTTAT 660
    |||||||
DB 13178 GTTATCGTTCTTAATGTTGACAGGACCAACGAGAGAGCGTATTGATACAGATATGTTAT 13119

QY 661 GATGACAAAGATTTATGTTTAAATGATCTCTGATGCGTAAAGAAAGAAAGATATAT 720
    |||||||
DB 13118 GATGACAAAGATTTATGTTTAAATGATCTCTGATGCGTAAAGAAAGAAAGATATAT 13059

QY 721 GAATCACTGCAAAATATTCAGTATTAAAGACCTTTAAAGCGATTGAACGTTCAATGTT 780
    |||||||
DB 13058 GAATCACTGCAAAATATTCAGTATTAAAGACCTTTAAAGCGATTGAACGTTCAATGTT 12999

QY 781 GTTTTACGTGTTATGATGACGAACAACGCAATTTTGAACAAGATAAACGTTGCACGA 840
    |||||||
DB 12998 GTTTTACGTGTTATGATGACGAACAACGCAATTTTGAACAAGATAAACGTTGCACGA 12939

QY 841 TATGACACATGAACAAGTAAAGACAGTCTGATTTGCTGAATTAATGGGATACGTGGAA 900
    |||||||
DB 12938 TATGACACATGAACAAGTAAAGACAGTCTGATTTGCTGAATTAATGGGATACGTGGAA 12879

QY 901 AAGATAGTAAACAGATGAACAATTTGAAGATGAGTAAAGATTTCCATTTTGA 960
    |||||||
DB 12878 AAGATAGTAAACAGATGAACAATTTGAAGATGAGTAAAGATTTCCATTTTGA 12819

QY 961 GATTATGCAAAATTTGCTTTGTGCTGCTAAGAACGACACAAAGTTACGATATATTC 1020
    |||||||
DB 12818 GATTATGCAAAATTTGCTTTGTGCTGCTAAGAACGACACAAAGTTACGATATATTC 12759

QY 1021 CCTTACATTAATGAAGACAGTGAACCAATTAACCAAGTTCCTCAACTTTAAT 1080
    |||||||
DB 12758 CCTTACATTAATGAAGACAGTGAACCAATTAACCAAGTTCCTCAACTTTAAT 12699

QY 1081 GAATGCTTACTGATGCAATTTCCATGACCTACACACAGACAAAGATAGACGTTTG 1140
    |||||||
DB 12698 GAATGCTTACTGATGCAATTTCCATGACCTACACACAGACAAAGATAGACGTTTG 12639

QY 1141 AATGCTTTTATGCAACAAGTTCATATGACACACGACATTTGTTATTTGTTAAT 1200
    |||||||
DB 12638 AATGCTTTTATGCAACAAGTTCATATGACACACGACATTTGTTATTTGTTAAT 12579

QY 1201 GATGTAAGATTAATGCAATTTTCTTATTAACGCTATTAGAGAAATCAATCCGTCGCT 1260
    |||||||
DB 12578 GATGTAAGATTAATGCAATTTTCTTATTAACGCTATTAGAGAAATCAATCCGTCGCT 12519

QY 12578 GATGTAAGATTAATGCAATTTTCTTATTAACGCTATTAGAGAAATCAATCCGTCGCT 12519
    |||||||
DB 12518 TTTGGTTTGAAGTACACCAATTCATTTATGATCGTCAAAAGAGA 12474

RESULT 2
AP003362/c 346900 bp DNA linear BCT 07-FEB-2002
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
sequence, section 5/9.
ACCESSION
AP003362 BA000017
VERSION
AP003362.2 GI:14247083
KEYWORDS
Staphylococcus aureus subsp. aureus Mu50
SOURCE
Staphylococcus aureus subsp. aureus Mu50
ORGANISM
Bacteria; Firmicutes; Bacillales; Staphylococcus.
```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source
1	Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I., Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M., Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A., Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C., Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J., Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C., Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.	Whole genome sequencing of methicillin-resistant <i>Staphylococcus aureus</i>	Lancet 357 (9264), 1225-1240 (2001)		Location/Qualifiers	1. 346900
					/organism="Staphylococcus aureus subsp. aureus Mu50"	gene
					/mol_type="genomic DNA"	CDS
					/strain="Mu50"	
					/sub_species="aureus Mu50"	
					/db_xref="taxon:158878"	
					12. .263	
					/gene="SAV1313"	gene
					12. .263	
					/codon_start=1	
					/transl_table=11	
					/product="hypothetical protein"	
					/protein_id="BAB57475.1"	
					/db_xref="GI:14247084"	
					/translation="MLTLIKLEDEQVILIEYIPDDISNGKSGVTFNKKAEDVDFS LSEINERYFMLYRNKSFVSVDFTLEKQEPENYKIAMY"	gene
					987. .1181	
					/gene="SAV1314"	CDS
					987. .1181	
					/gene="SAV1314"	gene
					/codon_start=1	
					/transl_table=11	
					/product="hypothetical protein"	
					/protein_id="BAB57476.1"	
					/db_xref="GI:14247085"	
					/translation="MNLQSYIGKLVKLTFTNNKILIGKVIDEDDKVDNFDGYSIEID TGRITYDISENKKIRIVLQK"	gene
					complement(1725. .2750)	
					/gene="SAV1315"	CDS
					complement(1725. .2750)	
					/gene="SAV1315"	gene
					/codon_start=1	
					/transl_table=11	
					/product="conserved hypothetical protein"	
					/protein_id="BAB57477.1"	
					/db_xref="GI:14247086"	
					/translation="MISFENDLEGAEHEKVLKRLVDTNLVQASGCFDFOFTQAIEKI KDIIDCPNATIRFELVGITQNOVAVINSMLSESEGISADTGVAHVEGALIEYSHK VITIPSEKRVASADIVETMETFSKDFKDHVFPQMVIYISHTEGTLYSKSELELC KVCQYQPLPMDGARLGLMDSDQSDMTIDIAKYCVFVIGTKIGALGCEALVFT KKNPKQFTTRIKHKGALLAGRLGTIOLELFTDNLVFNISRHAIEMANKKDKGFKN KGYRLYDPSPTNQOFTILSNEKIALBQKVFKAWEKTDQHRVVRPATSWATTEENL NKLEELT"	gene
					3020. .3217	
					/gene="SAV1316"	CDS
					3020. .3217	
					/gene="SAV1316"	gene
					/codon_start=1	
					/transl_table=11	
					/product="hypothetical protein"	
					/protein_id="BAB57478.1"	
					/db_xref="GI:14247087"	
					/translation="MEQIKTKFTFAETLELEKKNINAFLSSEATNLKLVNITKEIE ERFPNNEEPFNAIILLSVNK"	gene
					3238. .4755	
					/gene="SAV1317"	CDS
					3238. .4755	
					/gene="SAV1317"	gene
					/codon_start=1	
					/transl_table=11	
					/product="cardiolipin synthetase homolog"	
					/protein_id="BAB57479.1"	
					/db_xref="GI:14247088"	
					/translation="MEFYPPFIDGEKTMRYTFSNDIGTLFTIILAIGFIIINLVAFIII FLERNRRATSTWAMLEVLVPLIGFIIYLFEGRTVSARKLNKNGVLPDGLLK OOLSEPDKNGVGTDMKONKHNDIVSMILMDODGTLENNKVDHFDGNDLYDQVLDK IINAKETIHLEYTPRALDGLGRILHLEBEKIKOGIEYKILYDDVGSRYKVNANDHF KSLGGEVNEAFRSKILPLNFRNNRNHRKIIVIDQLGVGCFNIGDELGLGKLGIV RDTHLRIQGDVAVALQRLFLDMNSQAHRPQEPYDVKPPKRNGLPNSPIQIAASGP ASDWHQIEGYGKIMINSAKRSVYLQSPYIIPNSYINAIKIAKSGVDVHLMIPEKPD HDLVYVWATPSNMSDILSSGVKIYTYENGFIHSKMCILIDEIVSVGTANMPFSPELNF EVNAFYDENLAKDLRVANENDITTSKQLTRESYANRPLSVAFKESLAKLVSPIL"	gene
					4939. .5436	
					/gene="SAV1318"	CDS
					4939. .5436	
					/gene="SAV1318"	gene
					/note="ATP-binding protein"	
					/codon_start=1	
					/transl_table=11	
					/product="ABC transporter homolog"	
					/protein_id="BAB57480.1"	
					/db_xref="GI:14247089"	
					/translation="MIOISNINKSFNRKCVLKNISFDEIGKCIALIGKNGAKSTLI DILIGVANSGEIFDKDLQSENRSIMFOKMPDOLKVEIILNLOSFPENPLPL ERITFLTKDSSQLNQFVKLSGGORLLDFVLSLIGORQLLLDEPSTMDIETREY FMSIT"	gene
					5461. .5838	
					/gene="SAV1319"	CDS
					5461. .5838	
					/gene="SAV1319"	gene
					/note="ATP-binding protein"	
					/codon_start=1	
					/transl_table=11	
					/product="ABC transporter homolog"	
					/protein_id="BAB57481.1"	
					/db_xref="GI:14247090"	
					/translation="MLTSHYIEEVRMSDKIILLENGLIINDSTSHRTNQSQIT LSEYTRKLKLDKDDLVIOKNHNGTIITTSVNDOTILYLOOLHNLDDIETOKVSI VDSYFNKKORGSNYDTKLEENRI"	gene
					5807. .6538	
					/gene="SAV1320"	CDS
					5807. .6538	
					/gene="SAV1320"	gene
					/note="similar to ABC transporter integral membrane protein"	
					/codon_start=1	
					/transl_table=11	
					/product="hypothetical protein"	
					/protein_id="BAB57482.1"	
					/db_xref="GI:14247091"	
					/translation="MILSYLKEFRVIMRKKTLLISILFPVIEYILFTSILPEDEV KRFYKREVMYSWTVALSLSLTPPLIINKQEMQORLWNPFTFSYYSIVSVK TSLAIPALITLVFVGHGYKGVAMSAVOMLESGIFLMGASLLIFGLIFSLNDIQK TSLAINIYITGIAFGVGMPIINTPNNLQHAHVLPSYHLRKLKGLVDIASNNHNLIS FAIILYLGSTIAIAYCISHRRRAE"	gene
					6542. .7633	
					/gene="SAV1321"	CDS
					6542. .7633	
					/gene="SAV1321"	gene
					/note="similar to two-component sensor histidine kinase"	
					/codon_start=1	
					/transl_table=11	

```
/product="hypothetical protein"
/protein_id="BAB57483.1"
/db_xref="GI:14247092"
/translation="MKFLKDTSAIEISIIYLIFPIAGIFENVEYVGGKMLIISVYIF
SLSYLIVYNNKNTLMETMEYILILIHFFITCYEVSHPMLSEFFISARAVPTFEK
NVKLTATNLEILTMICTITTYLLYNNFYAMMYVYVILMLDNKKNNREYQKE
IAKNRHINTLIAEOERHRIQODLHGFASLSKSELAUKLIDTDEKVAELL
AINKLSRESINREIIDVDKLPSEIIDSIRKLDADIDFEEKKELAOVLSPTK
OSMLVMTTREATINNVIKHAANSKVHGKLVKNHKKLLMLIEDDGKIDSDCEVKSISO
RNOHNGTLVADSTNGKTIIEISTGIA"
7630..8232
/gene="SAV1322"
7630..8232
/gene="SAV1322"
/feature="similar to two-component response regulator"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="MTSLIAEDONMLRQAMVOLIKLHGDEILADPDNGIDAMKLE
EYRNVIYILDEMPKGLIEVLAIEIRKKHINIKIYITTKRREGYPERKAVYNDAYV
LKERSIEELVETINKVNGEKESATLMTSEYDKNPLTPEQOVLKNEIGLSSKEI
SEKLFITDGTIVRNTSVIIDLFPADNRFDMKKRNEKGMT"
complement(8345..8533)
/gene="SAV1323"
complement(8345..8533)
/gene="SAV1323"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNTSMENTVLLVIAIIIGIFLQIFKPLIVTAVLSILLGI
FVGRIYLIYSFNKRKN"
8672..9205
/gene="nuc"
8672..9205
/gene="nuc"
/feature="SAV1324"
/codon_start=1
/transl_table=1
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"

Query Match 99.08; Score 1292.2; DB 1; Length 346900;
Best Local Similarity 99.48; Pred. No. 1.2e-214;
Matches 1297; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

gene
CDS
1 ATGACTAACTATAGTATGTAGTGGCTATAGTAAATCTACAAATTTT 60
200810 ATGACTAACTATAGTATGTAGTGGCTATAGTAAATCTACAAATTTT 200751
61 AATGAGATAGTTGGAGACGTTTGCATTTGTAAGATACGCGAGGTGAACAGAGAT 120
200750 AATGAGATAGTTGGAGACGTTTGCATTTGTAAGATACGCGAGGTGAACAGAGAT 200691
121 CGTATTTATTTTCAAGTGAAGTGAACATGATTTCAATATTTATGATACAGTGTG 180
200690 CGTATTTATTTTCAAGTGAAGTGAACATGATTTCAATATTTATGATACAGTGTG 200631
181 ATTGAATTTGTTGATGACCAATTCACAAATTTAGAGCGAGCGAATATGCCCAT 240
200630 ATTGAATTTGTTGATGACCAATTCACAAATTTAGAGCGAGCGAATATGCCCAT 200571
241 GATGAAGCGATGTTATTTATTTATGTTAAGTGCCTGAAGATTTGACACAAAGCGAT 300
200570 GATGAAGCGATGTTATTTATTTATGTTAAGTGCCTGAAGATTTGACACAAAGCGAT 200511
301 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAACCGGTCGATTTAGCGGTTAACAA 360
200510 GAAATGTCGCTCAAAATTTTATACAAATCTAAAAACCGGTCGATTTAGCGGTTAACAA 200451
```

```
QY 361 GTATATATATGAAATGCGTACAGAGCTGATGATTTCTATTCATTAGATTTGGTGA 420
DB 200450 GTATATATATGAAATGCGTACAGAGCTGATGATTTCTATTCATTAGATTTGGTGA 200391
QY 421 CCGTATCCAAATATCAGAGTGCATAGTTTGGTGTGACTGTTAGATGACGTTGTT 480
DB 200390 CCGTATCCAAATATCAGAGTGCATAGTTTGGTGTGACTGTTAGATGACGTTGTT 200331
QY 481 TCTCATTTTGGTGAAGAGAGATCCCTATGATGAGTATGACAAATTCGCTATTCATT 540
DB 200330 TCTCATTTTGGTGAAGAGAGATCCCTATGATGAGTATGACAAATTCGCTATTCATT 200271
QY 541 ATTGACGACCAACAGTATGATTAATCAAGTTTATTAATGTTAGTGAAGATCGC 600
DB 200270 ATTGACGACCAACAGTATGATTAATCAAGTTTATTAATGTTAGTGAAGATCGC 200211
QY 601 GTTATCGTTTCTAATGTTGACGAGCAACGAGACGCTATTGATACAGACTATATGTTAT 660
DB 200210 GTTATCGTTTCTAATGTTGACGAGCAACGAGACGCTATTGATACAGACTATATGTTAT 200151
QY 661 GATGACAAAGTATATGTTTAAATGATACCTGCTATGCTTAAAGAAAGAAAGTATAT 720
DB 200150 GATGACAAAGTATATGTTTAAATGATACCTGCTATGCTTAAAGAAAGAAAGTATAT 200091
QY 721 GAATCAACTGAGAAATATTCAGTATTAAGACCTTTAAAGCATTTGAACTTCAATGTT 780
DB 200090 GAATCAACTGAGAAATATTCAGTATTAAGACCTTTAAAGCATTTGAACTTCAATGTT 200031
QY 781 GTTATGAGTATGATGACGAGAACAGCATTAATGAACAGATTAACGTTGGCAGCA 840
DB 200030 GTTATGAGTATGATGACGAGAACAGCATTAATGAACAGATTAACGTTGGCAGCA 199971
QY 841 TATGCACATGAAACAAGTAAAGCAGCTGCTATGCTGTAATTAATGGGATACGTGGAA 900
DB 199970 TATGCACATGAAACAAGTAAAGCAGCTGCTATGCTGTAATTAATGGGATACGTGGAA 199911
QY 901 AAGATAGTAAAGATGAAGAAATTTGAAGATGAGTACGTTAAAGATTTCAATTTTGA 960
DB 199910 AAGATAGTAAAGATGAAGAAATTTGAAGATGAGTACGTTAAAGATTTCAATTTTGA 199851
QY 961 GATTATGCAAAATTTGTTTGTGCTGCTAAGAACGACACAGATTAAGTATATTC 1020
DB 199850 GATTATGCAAAATTTGTTTGTGCTGCTAAGAACGACACAGATTAAGTATATTC 199791
QY 1021 CCTTACATTAATGAAGCAAGTGAAGAAACCATTAAGAAACGTTCAACATTTAAT 1080
DB 199790 CCTTACATTAATGAAGCAAGTGAAGAAACCATTAAGAAACGTTCAACATTTAAT 199731
QY 1081 GAAGTCGTACTGATGCAATTTCCATGAACCTACACCAACAGACAAAGATGAGAGTTG 1140
DB 199730 GAAGTCGTACTGATGCAATTTCCATGAACCTACACCAACAGACAAAGATGAGAGTTG 199671
QY 1201 GATGAGAAATTAATTCATTTTCTATTAAGGCTATTAAGGAATCAATCCGCGCGCT 1260
DB 199610 GATGAGAAATTAATTCATTTTCTATTAAGGCTATTAAGGAATCAATCCGCGCGCT 199551
QY 1261 TTTGGTTTGAAGTACACCAATTCATTTATAGTCGAAAGAGA 1305
DB 199550 TTTGGTTTGAAGTACACCAATTCATTTATAGTCGAAAGAGA 199506

RESULT 3
AX622668 1308 bp. DNA linear PAT 20-FEB-2003
LOCUS AX622668
DEFINITION Sequence 5631 from Patent WO02094868.
ACCESSION AX622668
VERSION AX622668.1 GI:28450653
KEYWORDS
```



```

SOURCE      Staphylococcus aureus
ORGANISM    Staphylococcus aureus
REFERENCE   1
AUTHORS     Masignani V.C., Mora M.C. and Scarselli M.C.
TITLE       Staphylococcus aureus proteins and nucleic acids
JOURNAL     Patent: WO 02094868-A 5631 28-NOV-2002;
            Chilton Spa (IT)
FEATURES    Location/Qualifiers
            source
            1..1308
            /organism="Staphylococcus aureus"
            /mol_type="genomic DNA"
            /db_xref="taxon:1280"
BASE COUNT  450 a 184 c 278 g 396 t
ORIGIN
Query Match      98.9%; Score 1290.6; DB 6; Length 1308;
Best Local Similarity 99.3%; Pred. No. 9.9e-214;
Matches 1296; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGTACTATTTGATAGGAGCCCTAATGTAATCTACAATTTT
    |||||
Db 1 ATGACTAAACCTATAGTACTATTTGATAGGAGCCCTAATGTAATCTACAATTTT
    |||||
QY 61 AATGACATAGTTGGAGACGCTGTTGCGATTTGGAAGATACGCCAGCTAACACGAGAT
    |||||
Db 61 AATGACATAGTTGGAGACGCTGTTGCGATTTGGAAGATACGCCAGCTAACACGAGAT
    |||||
QY 121 CGTATTTATTTCTTCAGGTGAGTGGTTAACACATGATTTCAATATTATTCACAGGTGT
    |||||
Db 121 CGTATTTATTTCTTCAGGTGAGTGGTTAACACATGATTTCAATATTATTCACAGGTGT
    |||||
QY 181 ATTTGAAATTTGGATGACACCATTTCCAAACAAATTTAGAGCCAGCAGAAATCCGCATA
    |||||
Db 181 ATTTGAAATTTGGATGACACCATTTCCAAACAAATTTAGAGCCAGCAGCAGAAATCCGCATA
    |||||
QY 241 GATGAAGCGGATGTTATTTTATGTTAAACGTCGCGTGAAGATTTGACACAAACGAT
    |||||
Db 241 GATGAAGCGGATGTTATTTTATGTTAAACGTCGCGTGAAGATTTGACACAAACGAT
    |||||
QY 301 GAAATGTCGCTCAATTTTATACAAATCTAAACAAACCGGTGCTATTAGCGCTTAACAA
    |||||
Db 301 GAAATGTCGCTCAATTTTATACAAATCTAAACAAACCGGTGCTATTAGCGCTTAACAA
    |||||
QY 361 GTAGATTAATATGAAATCGGTACAGACGCTGATGATTTCTATTCTATTAGATTTGGTGA
    |||||
Db 361 GTAGATTAATATGAAATCGGTACAGACGCTGATGATTTCTATTCTATTAGATTTGGTGA
    |||||
QY 421 CCGTATCCAAATACAGGGTCACATGGTTTAAAGTCTTGGTACTGTTAGATCAGTTGTT
    |||||
Db 421 CCGTATCCAAATACAGGGTCACATGGTTTAAAGTCTTGGTACTGTTAGATCAGTTGTT
    |||||
QY 481 TCTCATTTTGTGAAGAGAGAGATCCTTATGATGAAGATACAAATGCTACTATCCAT
    |||||
Db 481 TCTCATTTTGTGAAGAGAGAGATCCTTATGATGAAGATACAAATGCTACTATCCAT
    |||||
QY 541 ATTGACGACCAAAAGCTAGTAATCAAGTTTATGTAATGCTATTATTAGTGAAGATCCG
    |||||
Db 541 ATTGACGACCAAAAGCTAGTAATCAAGTTTATGTAATGCTATTATTAGTGAAGATCCG
    |||||
QY 601 GTTATCGTTTCTTAATGTTGCGAGGACACAGACGCTATTGATACAGATATAGTAT
    |||||
Db 601 GTTATCGTTTCTTAATGTTGCGAGGACACAGACGCTATTGATACAGATATAGTAT
    |||||
QY 661 GATGACACAGATTATGTTTATTCATCTACTGCTGATGCGTAAAAAAGAAAGATAT
    |||||
Db 661 GATGACACAGATTATGTTTATTCATCTACTGCTGATGCGTAAAAAAGAAAGATAT
    |||||
QY 721 GAATCACTGAGAAATATTTCACTATTAAGAGCTTTAAAAAGCATTTGAACGTTCAATGTT
    |||||
Db 721 GAATCACTGAGAAATATTTCACTATTAAGAGCTTTAAAAAGCATTTGAACGTTCAATGTT
    |||||
QY 781 GTTTTAGTGGTTATTGATGACAGACAAAGCATTTATTGAACAAGATTAACGTTGTGCAGGA
    |||||

```

```

Db 781 GTTTTAGTGGTATATGATCAGACAAAGCATTCATTGAACAAGATTAACGTTGTGCAGGA
    |||||
QY 841 TATGCACATGAAACAAAGCTAAGACGTCGATTTGCTGTAATTAATGGATCTGTGGA
    |||||
Db 841 TATGCACATGAAACAAAGCTAAGACGTCGATTTGCTGTAATTAATGGATCTGTGGA
    |||||
QY 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTCGAATTTTA
    |||||
Db 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTCGAATTTTA
    |||||
QY 961 GATTATGACAAATTTGCTTTTGTGCTGCTAAAGAACGACAGATTAAGTACATTAATC
    |||||
Db 961 GATTATGACAAATTTGCTTTTGTGCTGCTAAAGAACGACAGATTAAGTACATTAATC
    |||||
QY 1021 CCTTACATTAATGAAGCAAGTGAACCAATTAACCAATTAACCAATTAACCAATTAAT
    |||||
Db 1021 CCTTACATTAATGAAGCAAGTGAACCAATTAACCAATTAACCAATTAACCAATTAAT
    |||||
QY 1081 GAAGTCGCTACTGATGCAATTTCCATGAAACCCCTACACCAACAAAGTACGTTTG
    |||||
Db 1081 GAAGTCGCTACTGATGCAATTTCCATGAAACCCCTACACCAACAAAGTACGTTTG
    |||||
QY 1141 AATGCTTTTATGCAACACAGTGTCTATGAAACACGACATTTGTTGTTATTTAT
    |||||
Db 1141 AATGCTTTTATGCAACACAGTGTCTATGAAACACGACATTTGTTGTTATTTAT
    |||||
QY 1201 GATGTAGATTAATGATGATTTTCTTATTAACGCTATTGAGAAATCAATCCGTCGCGT
    |||||
Db 1201 GATGTAGATTAATGATGATTTTCTTATTAACGCTATTGAGAAATCAATCCGTCGCGT
    |||||
QY 1261 TTTGGTTTGAAGGTACACCAATTCATATTATAGTCGGAAGAAGA 1305
    |||||
Db 1261 TTTGGTTTGAAGGTACACCAATTCATATTATAGTCGGAAGAAGA 1305
    |||||
RESULT 4
AP004827/c 333750 bp DNA linear BCT 24-APR-2003
LOCUS      Staphylococcus aureus subsp. aureus MM2 complete genome,
DEFINITION strain:MM2, section 6/10.
ACCESSION AP004827 BA000033
VERSION   AP004827.1 GI:21204509
KEYWORDS  Staphylococcus aureus subsp. aureus MM2
SOURCE    Staphylococcus aureus subsp. aureus MM2
ORGANISM  Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
            Genome and virulence determinants of high virulence
            community-acquired MRSA
            Lancet 359 (9320), 1819-1827 (2002)
TITLE     Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
            Kikuchi,H.
            Direct Submision
            Submitted (06-MAR-2002) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center, 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, url:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES  Location/Qualifiers
            source
            1..333750
            /organism="Staphylococcus aureus subsp. aureus MM2"
            /mol_type="genomic DNA"
            /strain="MM2"
            /db_xref="taxon:196620"
            complement(278..619)
gene

```



```

CDS      /gene="WM1341"
          complement(278..619)
          /gene="WM1341"
          /note="ORFID:WM1341"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAB95206.1"
          /db_xref="GI:21204510"
          /translation="MTNITKEVDNLBOEIOVEAKNKTLSGSEAKPYLDEHYSKIID
          EKVYNDITGNIDFDKLNQYPPVPMNFKERYDYMIRKTHFMGYROMKTFKELLIMNA
          STQRLKKNQV"
gene      complement(624..1283)
          /gene="hth"
          /note="hth"
          /codon_start=1
          /transl_table=11
          /product="endonuclease-like protein"
          /protein_id="BAB95207.1"
          /db_xref="GI:21204511"
          /translation="MWSKKALEMIDVIANMPPDAECCELDKNDNPELTITAVLSAOCT
          DVLNRYVTELEFKYKTPEDYLAVSDEELMNDIRSIGLRNKAQIKKLCOSLIDQFN
          GEIPQTHKELESAGVGKRTANVMSVAFDEPSLAVDTHVERVSRLGINRMKDNVRQ
          VEDRLCSVIPRDMRMRSHHQLIFGGRYHCLARKPKCDICPLEDCREGOKRYKASLKE
          A"
gene      complement(1273..1959)
          /gene="WM1343"
          /note="WM1343"
          /note="ORFID:WM1343"
          /protein_id="BAB95208.1"
          /db_xref="GI:21204512"
          /translation="MDKYOLKARPVIRRELLDHYSDIGLDEODLVILLKIYASETS
          NKPSIELLOKSGTMQRPDITMVTONLQRELELOVOKDEGRTEVNNIDPFPEKL
          SHILKQSMETREONSKKFKQLFVLDEDTFARPLSPYELETLNOMIVDVKHDSAIIO
          AALDEANLNLSEFKYMDRILLNMKKNVKTIDSRKIREKFNKPKMTHYKTVKFD
          WLNENLNDGK"
gene      complement(2287..3579)
          /gene="asns"
          /note="asns"
          /complement(2287..3579)
          /gene="asns"
          /note="ORFID:WM1344"
          /codon_start=1
          /transl_table=11
          /product="asparaginyl-tRNA synthetase"
          /protein_id="BAB95209.1"
          /db_xref="GI:21204513"
          /translation="WKTTIKQAKDHLNODVTIGAWLTNRKSSGKIAPLQLRDGTGFMQ
          GVYVASEVDEEYFKLAKETIOESSLYVTGITEDNRSDLGEMQKSIIEVSEADPY
          ITPKNHGTLEFMDHLRLRSKQCHAVMKIRNEVIRATYEFENKDGFTKVDPILTAS
          APEGSELFTHTKTFQDAFLVOSGOLYLEAAMAHGKVFSGPTFRRAKSKRRRLIE
          FMMEGEMAFNHAESLEFIOQOYTHVVKYLVENKELKILERTPSLKEKVAATPPR
          ISYDDAIFLKSSEGDDILEGSEDGAPVHETAIANNHIDLPVFTTNPYKIKPPYQPNP
          ENETVLCADLIAPGYGELIGGSEKRVDELLEORVKEHGDSEAVSYVLDLRRYGS
          VPHGFGGLGERTYAMISGVEHVRETAFPRILNRLYP"
gene      complement(3901..6594)
          /gene="ding"
          /note="ding"
          /codon_start=1
          /transl_table=11
          /product="probable ATP-dependent DNA helicase ding"
          /protein_id="BAB95210.1"
          /db_xref="GI:21204514"
          /translation="MGATYAVVDDLETTGNQLDPDDDIIOIGITFVRNNOIIDTYHSMI
          RTNLIPPIQALTSIEENMLQOAPYFNQVAOETDYKIKDCIFVAHNVDFLNFIKKA

```

```

gene      complement(6618..7589)
          /gene="WM1346"
          /note="WM1346"
          /codon_start=1
          /transl_table=11
          /product="truncated hypothetical protein"
          /protein_id="BAB95211.1"
          /db_xref="GI:21204515"
          /translation="MSKYSODVLAQLYKKNKPNYISGQSIASLSNTPYAVKVIDQKL
          LEGCIDSVDNHRGHLLQOLPDWOGIIDQYTKSSALPDFSEKVDSDISTOLAKKSL
          VGNQSSFILSDQETKGRGRFRNRHSSKCGQGLMMSVYLRPVAVASMSKRNLTALG
          IRDAIOHFSODEKVKMPNDYITDNGKCGEFTREVANNNDGEALICGIGLITQOLE
          NFDESIRHRATNSVOLHDKNKLDRQVLELQIEKRVNQLTFPSEIREYIYASN
          IWNRTILFTENDQFKGOAIDLDYDGYVLRVDEAGSHRLISADIDF"
gene      complement(7576..8778)
          /gene="WM1347"
          /note="WM1347"
          /note="ORFID:WM1347"
          /protein_id="BAB95212.1"
          /db_xref="GI:21204516"
          /translation="MDKSLFEDQAPILBOIODNGFEAVYVGGSVRDYVWGRNIDIDI
          TTSATPDEIESIFSHTIPEGKHGTINVFENDENEVTFRABEDYVDHRRSQTFFV
          RDLVDELQRODFTNNAIAMDYAKLYDYDFGQODINNHIRTVGIAERFQEDALRMI
          RCLRFOSQSFQIAETFEFAMRIQADIKFLSIEIVIELTKLMGIVNEKSFNLKS
          LKAFVMPYFELDMNOJNVEAIDLELLIAVSVPDINYSIKLXKISNRQVNDIO
          YIQINNALPSITTKSQLKNFYVDYTHILIKNVAADVIKANDIGKHEPLVYNLOTTD
          ETLHRLPMMNRKDMVNGCVLMAHLNAKSGPWLKQVLOIEIAIVTGKVSNEETELK
          WVDNHVKI"
gene      complement(8783..9925)
          /gene="WM1348"
          /note="WM1348"
          /note="ORFID:WM1348"
          /protein_id="BAB95213.1"
          /db_xref="GI:21204517"
          /translation="MKIGITCYPSMGSGQIATELGIKLAERGHEVHTTSNIPRIR
          KPLPMIFHOVNOYAVEQYPPDYITLSTKIAEVIKEYGDLDMHYAPHAIGIL
          AREMSGDKIKMTLTGHDITIVIGYDHSGLQAKIKGKISDSDVTSVSKSLAOEHTIIT
          ETNKEITIIYVNVRENERTPKNTALAKQFGIAPDEKVLINHSNROVKRRIITIEF
          AKVREKIPSKLILGDGPELVPKROLIRLELNEEDVLELGKDCYSEFYQUSDVLL
          SEKESFGLLEAMKTVGVPISGNSAGIKEYIKHGETGFVNDVGDQSDASDAIRLLE
          DKVLNKLQKNMLADIARFGSELITTDQYEVYYQMLNEHNSKGE"
gene      complement(10167..10484)
          /gene="WM1349"
          /note="WM1349"
          /codon_start=1
          /transl_table=11

```

/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translation="MKSVMQREVDVYIGQFGYFSPLANLARLEEVGELAREIN
HTYGEKKKKDSEBANTIKAEIDGNLFVLLCLANSMGIDMTESNEMTEKMFNTDKRRF
EKK"
complement(10820..11518)
/gene="MM1350"
complement(10820..11518)
/gene="MM1350"
/note="ORFID:MM1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translation="MSSIFMTIIEVILWLPLMAQHKVSNKEYSOVRSYSGTKGR
EVALIELHANGIYDVEVVKGEFLTDHYOPKKVYLSLSPANDRPSVAATLAAHEVG
HAIQDHQWELFRALVFNANLSSLSLTIILITLALISAFSGTALMIGAGLM

Query Match 98.9%; Score 1290.6; DB 1: Length 333750;
Best Local Similarity 99.3%; Pred. No. 2,3e-214;
Matches 1296; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGACTAAACCTATAGTATGATGAGGCTTAATGTAGCTAAATCTACAAATTTT 60
DB 24351 ATGACTAAACCTATAGTATGATGAGGCTTAATGTAGCTAAATCTACAAATTTT 24292
QY 61 AATGAGATGATGAGAAAGCTGTTGATGTTGAAGATACGCGAGGTGAACAGCAT 120
DB 24291 AATGAGATGATGAGAAAGCTGTTGATGTTGAAGATACGCGAGGTGAACAGCAT 24232
QY 121 CGTATTATCTTCAGGTAGTGTAGACATGATTTCAATATATTATAGAGGTGT 180
DB 24231 CGTATTATCTTCAGGTAGTGTAGACATGATTTCAATATATTATAGAGGTGT 24172
QY 181 ATTGAATGTTGTATGCACCATTCCTCAACCAAAATTAGACCGACAGCAATCCCAT 240
DB 24171 ATTGAATGTTGTATGCACCATTCCTCAACCAAAATTAGACCGACAGCAATCCCAT 24112
QY 241 GATGAAGCGGATGTTATATTTTATGTTAAACGTCGTAAGGATTTGACAAAGCAT 300
DB 24111 GATGAAGCGGATGTTATATTTTATGTTAAACGTCGTAAGGATTTGACAAAGCAT 24052
QY 301 GAAATGTCGCTCAAAATTTTATCAAAATCTAAAAAACCGGTCGTATTAGCGGTTAA 360
DB 24051 GAAATGTCGCTCAAAATTTTATCAAAATCTAAAAAACCGGTCGTATTAGCGGTTAA 23992
QY 361 GTAGATAATATGAAATGCGGTACAGACGCTATGATTTCTATTCTATTAGATTTGGTGA 420
DB 23991 GTAGATAATATGAAATGCGGTACAGACGCTATGATTTCTATTCTATTAGATTTGGTGA 23932
QY 421 CCGTATCCAAATATCAGGTCACATGTTAGTCTTGCGATCTGTTAGATGACAGTTGT 480
DB 23931 CCGTATCCGATATCAGGTCACATGTTAGTCTTGCGATCTGTTAGATGACAGTTGT 23872
QY 481 TCTCATTTTGGTGAAGAGAAGATCCTTATGATGAAGATCAATTCGACTATCCATT 540
DB 23871 TCTCATTTTGGTGAAGAGAAGATCCTTATGATGAAGATCAATTCGACTATCCATT 23812
QY 541 ATTGAGCAACAAGCTAGGTAATCAAGTTTGTAAATCCCTTTTATAGTGAAGATCCG 600
DB 23811 ATTGAGCAACAAGCTAGGTAATCAAGTTTGTAAATCCCTTTTATAGTGAAGATCCG 23752
QY 601 GTTATCGTTTCTAATGTTCAGGAGACACAGAGCGCTATTATACAGAGTATAGTTAT 660
DB 23751 GTTATCGTTTCTAATGTTCAGGAGACACAGAGCGCTATTATACAGAGTATAGTTAT 23692
QY 661 GATGACAAGATTATGTTTAAATCGATCTGCTGATGCGGTAAAAAGAAAAAGTATAT 720
DB 23691 GATGACAAGATTATGTTTAAATCGATCTGCTGATGCGGTAAAAAGAAAAAGTATAT 23632
QY 721 GAATCAACGAGAAATATTCAGTATTAAGAGCTTTAAAGCGATTGAACGTTCAAAATGTT 780

DB 23631 GAATCAACGAGAAATATTCAGTATTAAGAGCTTTAAAGCGATTGAACGTTCAAAATGTT 23572
QY 781 GTTATCGTTTCTAATGTTCAGGAGACACAGAGCGCTATTATACAGAGTATAGTTAT 840
DB 23571 GTTATCGTTTCTAATGTTCAGGAGACACAGAGCGCTATTATACAGAGTATAGTTAT 23512
QY 841 TATGACATGAAACAGGTAAGAGCATGCTGATTTGCTGAATTAATGATGATCTGTGAA 900
DB 23511 TATGACATGAAACAGGTAAGAGCATGCTGATTTGCTGAATTAATGATGATCTGTGAA 23452
QY 901 AAGATAGTAAACGATGAAAGAAATTTGAAGATGAAAGTACGTTAAAGATTTCAATTTTA 960
DB 23451 AAGATAGTAAACGATGAAAGAAATTTGAAGATGAAAGTACGTTAAAGATTTCAATTTTA 23392
QY 961 GATTATGACAAATTCCTTTTGTGCTGCTGAAGAGCACAAGATTAGCTACATTATTC 1020
DB 23391 GATTATGACAAATTCCTTTTGTGCTGCTGAAGAGCACAAGATTAGCTACATTATTC 23332
QY 1021 CCTTACATTAATGAAACAGTGAAGAAACCATTAAGAAAGCTGTTCAAGTCACTTTAAAT 1080
DB 23331 CCTTACATTAATGAAACAGTGAAGAAACCATTAAGAAAGCTGTTCAAGTCACTTTAAAT 23272
QY 1081 GAAGTGTATCTGATGCAATTTCCATGAAACCCCTACACACAGCAAGGTAGACGTTTG 1140
DB 23271 GAAGTGTATCTGATGCAATTTCCATGAAACCCCTACACACAGCAAGGTAGACGTTTG 23212
QY 1141 AATGCTTTTATGCAACACAGTGTATGAAACCCGACGATTTGTTGATTTGTTAAT 1200
DB 23211 AATGCTTTTATGCAACACAGTGTATGAAACCCGACGATTTGTTGATTTGTTAAT 23152
QY 1201 GATGTGAATTAATGCAATTTTCTTATTAACGCTATTATGAGATCAATCCGTGCCCT 1260
DB 23151 GATGTGAATTAATGCAATTTTCTTATTAACGCTATTATGAGATCAATCCGTGCCCT 23092
QY 1261 TTGCTTTGAAAGTACACCAATTCATATTATGCTGGAAGAGA 1305
DB 23091 TTGCTTTGAAAGTACACCAATTCATATTATGCTGGAAGAGA 23047

RESULT 5
AF270032/c 3269 bp DNA linear BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1042f07 genomic
DEFINITION sequence.
ACCESSION AF270032
VERSION AF270032.1 GI:9623936
KEYWORDS
SOURCE Staphylococcus epidermidis
ORGANISM Staphylococcus epidermidis
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 3269)
AUTHORS Kimmerly,W.J., Taylor,J. David., Nelson,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenebee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3269)
AUTHORS Taylor,J. David., Kimmerly,W.J., Nelson,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenebee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source 1..3269
/organism="Staphylococcus epidermidis"

/mol_type="genomic DNA"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1042f07"
BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN

Query Match 70.5%; Score 919.4; DB 1; Length 3269;
Best Local Similarity 81.5%; Pred. No. 1.3e-149;
Matches 1064; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

OY 1 ATGACCTAAACCTATATAGTATTTGTAGTAGGCGCTAAATGTAGTAATCTACAATTTT 60
DB 2804 ATGACTAAACCGATAGTACAAATTTGTAGAAAACCAATGTAGAAAATCTCAATTTT 2745
OY 61 AATAAATAGTTGGAGAACGTTTCGATTTGTGAAGATACCCAGGTGTACAGAGAT 120
DB 2744 AATAAGTTGTGCGGAACGTATATCTATCGTAGAAGATACGCCAGGCGTTCTCGTGAT 2685
OY 121 CGTATTTATTCCTTCAGGTGAGTGGTTAACACATGTTCAATATTTATTTGATACAGGTGT 180
DB 2684 CGTATTTATTCATCTGGGAGAGGTGTTAACTCATGATTTTACATCATTCATACAGGTGT 2625
OY 181 ATTGAATTTGGTATGCACCATTTCCAAACACAAATTAGAGCGCAGGACAGAAATCCGCATA 240
DB 2624 ATTGAATTTGGAGATGCTCTCTTTTCAAACTCAAAATCCGTCAGACAGGACAGAAATAGCAATT 2565
OY 241 GATGAAGCGGATGTATATTTTATTTGTATACGTCGCGTGAAGGATGTACACAAAGCGAT 300
DB 2564 GAAGAAGACAGATGTATCTATTTTATGTGCAATTTAGAGAAAGGACTTACACAAAGTGAC 2505
OY 301 GAATAGTGTGCTCAATTTTATACAAATCTAAAAACCGTGTATTTAGCGGTTAACAA 360
DB 2504 GAATAGTGTGCGCAAAATGCTTATTAATCTAAGAAAACCGTGTATTTAGCTGTGAATTA 2445
OY 361 GTAGTAATTAATGGAATGCGTACAGACGTTGATGATTTCTATTTCAATTTAGATTTGGTGA 420
DB 2444 GTTGATTAATCTGAATGCGTAAATGATATCTATGATTTCTATTTAGCTTTGGAGAT 2385
OY 421 CCGTATTCGAATTCAGGCGTACATGTTAGTCTTTGGGACTTTGTAGATGCACTGTT 480
DB 2384 CCATATCTATTTTGTGTTACATGATTAAGTGAAGTGTGAGATTTCTATGATTCAGCTGTT 2325
OY 481 TCTCATTTTGTGGAAGAGAAAGATATCTTATGATGAAGATACAAATTCGATTCATC 540
DB 2324 GAAACCTTTAATTAAGAAATCAGAAATCTTTATGACGAAGATACGATCTTTCTATAC 2265
OY 541 ATTGACGACCAAAACGTAGTAAATCAAGTTAGTAATGCTATTTTAGTAAAGATACGC 600
DB 2264 ATCGGTAGACCTAATGTTGGTAAATCTGACCTTGGTCAATGCTATTTTAGCGAAGACGT 2205
OY 601 GTTATCGTTTCTAATGTTGACAGGACAAAGACAGACGCTATTGATACAGATATGTTAT 660
DB 2204 GTTATTTGTCTAAATGTTGCTGCTACAACTCGAGATGCCATTCACCACTGCTCTTAT 2145
OY 661 GATGACAAAGATTTATGTTTATTCGATCTGCTGATGCGTAAAAAGAAAGATATAT 720
DB 2144 GATGACAAAGATTTATGTTTATGTTGATCTGCTGATGAGAAAGAAAGATAGGTGTTAT 2085
OY 721 GAATCACTGAGAAATATTCAGTATTAAGACCTTTAAAGCGATTGAACGCTCAAAATGTT 780
DB 2084 GAATGACCTGAGAAATATTCCTGATTAAGCTGATTTAAAGCGATTGAGGCTTCAAGATTA 2025
OY 781 GTTTTAGTGTATTTGATGACAGACAGGACATTTTGAACAAGATTAACGCTGTTGACAGA 840
DB 2024 GTATTAGTGTATTTGATGACAGACAGGATTAATTAAGATTAACGCTGTTGACAGA 1965
OY 841 TATGACATGACAAAGGTAAAGCAGTGTGATTTGCTGAATTAATTAAGGATCTGCTGAA 900
DB 1964 TATGACATGACAAAGGTAAAGCAGTGTGATTTGCTGAATTAATTAAGGATCTGCTGAA 1905
OY 901 AAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTTCAATTTTAA 960

DB 1904 AAAGATAGTAAACGATGAAGAAATTTCACTGATGATTTAGAAATGTAATTTCAATTTTAA 1845
OY 961 GATTATGACCAAAATTTGCTTTGTCGCTTAAGAAACGCAAGATATTACGATATTATTC 1020
DB 1844 GATTATGCTCAAAATGCGCTTCTATACGAAAGAAAGAGGCTTAAGATTTAAACCATTTATTC 1785
OY 1021 CCTTACATTTAATGAAGCAAGTGAAGAAACCATTAAGAAACGTTTCAAAAGTTCAACTTTAAAT 1080
DB 1784 CCTTATATTCATCAACGCGAGTGAAGAAATCTAATAAAGCGTGTCCAAAGTTCTACAGCAAAAT 1725
OY 1081 GAATGCTTACTGATGCAATTTTCCATGAACCTTACACACAGACAGACAAAGTGTAGCGTTTG 1140
DB 1724 GAATGCTTACTGATGCAATTTTCCATGAATCCAAACCTTACAGACAAAGTGTAGAAAGACTT 1665
OY 1141 AATGCTTTATGCAACCAAGTGTGATAGAACACGCAATTTGTTGTTATTTGTTAT 1200
DB 1664 AATGCTTTATGCAACCTGATGTTGAATGAACCAACGCAATTTGTTGTTATTTGTTAT 1605
OY 1201 GATGTAAGATTAATGCAATTTTCTTATAAACGCTATTTTGAAGATCAAAATCCGTGCGCT 1260
DB 1604 GATGTTGATTAATGCAATTTTCTTATAGAGATATTTGAAATCAAAATACGATATGCT 1545
OY 1261 TTTGCTTTTGAAGTACACCAATTCATATTATAGCTGCAAGAGA 1305
DB 1544 TTTGCTTTTGAAGAACACCTATTCATATTATTTCCAAAGAAAAGA 1500

RESULT 6
AX145350/c 3269 bp DNA linear PAT 31-MAY-2001
LOCUS AX145350
DEFINITION Sequence 4072 from Patent WO0134809.
ACCESSION AX145350
VERSION AX145350.1 GI:14283915
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Kimmerly, W. J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4072 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source 1. 3269
location/Qualifiers
1. 3269
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"

BASE COUNT 986 a 616 c 451 g 1216 t
ORIGIN

Query Match 70.5%; Score 919.4; DB 6; Length 3269;
Best Local Similarity 81.5%; Pred. No. 1.3e-149;
Matches 1064; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

OY 1 ATGACCTAAACCTATATAGTATTTGTAGTAGGCGCTAAATGTAGTAATCTACAATTTT 60
DB 2804 ATGACTAAACCGATAGTACAAATTTGTAGAAAACCAATGTAGAAAATCTCAATTTT 2745
OY 61 AATAAATAGTTGGAGAACGTTTCGATTTGTGAAGATACCCAGGTGTACAGAGAT 120
DB 2744 AATAAGTTGTGCGGAACGTATATCTATGTAAGATACGCGACGCTTACTCGTGAT 2685
OY 121 CGTATTTATTCCTTCAGGTGAGTGGTTAACACATGATTTTAATTTATTTGATACAGTGT 180
DB 2684 CGTATTTATTCATCTGGGAGAGGTGTTAACTCATGATTTTAACTTATTTGATACAGTGT 2625
OY 181 ATTGAATTTGGATGATGACCATTTCCAAACACAAATTAAGAGCGGACAGAAATCCGCATA 240
DB 2624 ATTGAATTTGGAGATGCTCTTTTCAAACTCAAAATCCGTCGACAGGACAGAAATAGCAATT 2565
OY 241 GATGAAGCGGATGTTTATTTTATGTTTAAAGTGTGCGTGAAGATTTGACACAAAGCGAT 300

Db 2564 GAAGAAAGCATGTCATCATTTTATGCTCAATGTTAGAGAGAGACTTACACAAAGTGAC 2305
QY 301 GAAATGGTCGCTCAAAATTTTATACAAATCAAAAACCGGTCGATTAGCGGTTAACAAA 360
Db 2504 GAAATGGTCGCAAAATGCTTTTAAATATAGAGAAACCTGTTGATTAGCTGGAATAAA 2445
QY 361 GTAGATTAATATGAAATGCTTACAGAGCTATGATTCTTATTCATTAGCATTTGGTGAA 420
Db 2444 GTTGATTAATCTTAAATGCGTAATGATATGATTTCTTATCTTATAGCTTTGGAGAT 2385
QY 421 CCGTATCCATATACAGGTCACATGCTTGTAGCTCTTGCTGACTTGTGATGAGAGTGTT 480
Db 2384 CCAATTCCTATTTCTGCTTACATGCAATGAGCTTGGAGATTTGCTAGATTCAGTTGTT 2325
QY 481 TCTCATTTTGTGAGAGAGAGAGATCCTTATGATGAGATCAATTTGACATATCCATT 540
Db 2324 GAAACCTTTTAAATAGAAATCAGAAAGATCCTTATGACAGATATGAGATGCTTTTATC 2265
QY 541 ATTGAGAGCAACCAACGTAATATCAAGTTTATGTAATGCTTATTTAGTGAAGATCGC 600
Db 2264 ATGCGTAGACCTATATGTTGGTAATCTAGCTGCTCAATGCTATTTTATGCGGAGAACGT 2205
QY 601 GTTATGCTTTCTAATGCTTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 2204 GTTATGCTGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2145
QY 661 GATGACAGAGATTAATGTTTAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 2144 GATGACAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2085
QY 721 GATGACAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 2084 GATGACAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025
QY 781 GTTATGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 2024 GTTATGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1965
QY 841 TATGACAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 1964 TATGACAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1905
QY 901 AAGATAGTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1904 AAGATAGTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845
QY 961 GATTATGACAAATGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1844 GATTATGCTCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1785
QY 1021 CCTTACATTAATGACAG 1080
Db 1784 CCTTATATCAATCAACAG 1725
QY 1081 GAATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1724 GAATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
QY 1141 AATGCTTTTATGACAG 1200
Db 1664 AATGCTTTTATGACAG 1605
QY 1201 GATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1604 GATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1545
QY 1261 TTTGCTTTTGAAGTACCAATTCATTTATATAGCTGGAAGAGAGA 1305
Db 1544 TTTGCTTTTGAAGTACCAATTCATTTATATAGCTGGAAGAGAGA 1500

RESULT 7
AB016747/c

LOCUS AB016747 300892 bp DNA linear BCT 01-JAN-2003
DEFINITION Staphylococcus epidermidis ATCC 12228, section 4 of 9 of the complete genome.
ACCESSION AB016747
VERSION AB016747.1 GI:27315369
KEYWORDS Staphylococcus epidermidis ATCC 12228
SOURCE Staphylococcus epidermidis ATCC 12228
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE Zhang, Y., Ren, S., Li, H., Fu, G., Lu, L., Lu, G., Jia, J., Tu, Y., Qin, Z., Chen, Z. and Wen, Y.
Direct Submission
Submitted (05-NOV-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China
FEATURES
source
1..300892
/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:176280"
113..346
/locus_tag="SE0907"
113..346
/locus_tag="SE0907"
/codon_start=1
/transl_table=11
/product="acyl carrier protein"
/protein_id="AA004504.1"
/db_xref="GI:27315370"
/translation="MENFDKVKDIIIVRLGVADKVTEDASFQKDLGADSLDIAELVM
ELEDFEGTEIPDEAEKINIVGDAVYINSLEK"
461..1198
/locus_tag="SE0908"
461..1198
/locus_tag="SE0908"
/codon_start=1
/transl_table=11
/product="Rnase III"
/protein_id="AA004505.1"
/db_xref="GI:27315371"
/translation="MANOKKKEVNHPOOKFTDMKMSLGHFKFNIDLYOAPSHSFI
NDNMNRLEHNERLEFVGLVLELVSRILEPRLPBLNLTMRATTVCESEVIF
ANKIKLNEILILGKEGKEGTGRPSLIDAEFAFGALYLDGLDSVTFARVIFP
YVEDDELIVGVDFKTOFOEYVHSONKGDVYQLIKEGPAHRLTFSEVILEKAAVE
GKKTRESPOKABOAYKLMMKKSL"
1325..4894
/locus_tag="SE0909"
1325..4894
/locus_tag="SE0909"
/codon_start=1
/transl_table=11
/product="chromosome segregation SMC protein"
/protein_id="AA004506.1"
/db_xref="GI:27315372"
/translation="AVYLKSIDAIGFSPADHTNVOFDKGVTAIVPNSGKSNITDA
IKWILEGSAKSLRGAEMEDIFSGAEHKAQFAVEFKLNDHSQKLDIDSEELVT
RLYRSGSESYIYNDRARLKIDILDFDSGLGKEFSLISGRRVEILNARINRRO
IIFESAGVLYKKRKRAESTOKLDHPEFNDLQINRYREDILYDLEGREPKKEBATAKEKO
LSKEMQSDVITYSDIDHTEDNQLDLRLNLKSOQKEBQQAQIQLDLQKRGK
RQNDNDIEKLNLVAKATENYEQLSGKLVLEERKNQSEINARYEELDNLEQID
SINKERAQNEKLIAELKRNQOKLNEVQESLILYISDQHEKELEIKNSYTLMS
OSDYNNDIREFLEHTINENAKKSRDLRSLVEAFNQLDIQONITQOKEYOSKSM
KVQONIDQLEOULTDLSKRLISEVENKLYOAYRNEKLKRISLATOQEDDYTFENYV
KHILKAKDELKRGIRGAVAEVIVPSEMTQALITAGASLQHYIYDNEQDQAIOLY
KQGLDLRATFLPLNVLOPRHVADIKDVARSGQFTINISDAINVSATQONITENLGS
NTIIVENLKHANLARAIRYRTIIVLEGDVNPVPGSMGSGARKTSILISQDELST
MRNQLDYOROTAEFEFROPEKQTOAEQSLSEQYFSAQOYNNILKEQVHHLELDLTK
TOEHLNKEHEEPKNDGYOSDKETLKEKQNLHLEIIOOQLKQLEDIDERYTOLS
KEKASTHOTQOOLHOKQSLDAVYKRIKSOYVRIKQDLKQSLDSROKIEVKEIKL
FNSDEMKGDAFEKLEQIIOQDENVQONLNQDLSFKQQRKDLNEIETINESLQKCH
ODLSTENHYODIKAKOSKIDVLINLIDHLDNYTOLYVERARMEVDSQETIDNLKK

Db	297004	TTTGAATTTGGAGATGTCCTTTTCAAACTCAAAATCCGTCGACAGGACAGAAATPAGCAAT	296945
QY	241	GATGACGGCGATGTATTATTTTATGTTAACTGCGCTGAAGAGATTGACACAAACCGAT	300
Db	296944	GAAGACAGAGATGTCATCTATTTTATGTCATGTTAGCAAGGAGCTTACAAAGATGAC	296885
QY	301	GAATAGCGCGCAATTTTATATACAAATCTAAAAACCGGTCGATTTAGCGGTTAACAA	360
Db	296884	GAATATGTCGCACCAATGCTTTATATAACTTAAGAAACCTGTGTATTTACTGTGAATPAA	296825
QY	361	GTAGATTAATATGGAATAGCGTACAGACGATGTGATTTCTATTCAATTTAGACTTTGGTGA	420
Db	296824	GTTGATATATCTGAATAGCGTAAATGATATCTATGATTTCTATTCTTTAGCTTTGGAGAT	296765
QY	421	CCGTTATCCAAATACAGGGCTACATGCTTTAGGCTTTGGTACTTGTTAGATGCAGTTGTT	480
Db	296764	CCATATCTTATTTCTGGTTTCACATGATGATTAGGACTTTGGAGATTTTGGCTATGATCAGTTGTT	296705
QY	481	TCTCATTTTGGGAGAGAGAACAAATCCTTATGATATGAAAGATACAAATTCGACTATTCAT	540
Db	296704	GAAACTTTTAATTAAGAAATCAGACAAATCCTTATGACGAAGATACGATACGCTCTTCTATC	296645
QY	541	ATTGACAGCCAAACAGTAGTAAATCAAGTTTATGTAATGCTATTTTTATGAGTAAGATCGC	600
Db	296644	ATCGGTAGACCTTAATGTTGGTAAATCTAGCTTGTCGTAATGCTATTTTAGCGCAAGACGT	296585
QY	601	GTTATCGTTTCTAATGTGGCAGGCAACAGAGACGCTATTGATACAGATATAGTAT	660
Db	296584	GTTATATGTCCTTAATGTCTGGTGCACACTCGAGATGCTATGATACGAGATACCTTTAT	296525
QY	661	GATGACACAAGATTATGTTTATATCATCTCGTGGTATCCGTAATGAAAAAGAAAGATATAT	720
Db	296524	GATGACACAAGATTATGTTATGATGATGATCTGCTGGAATGAGAAAAAGTAAGGTAT	296465
QY	721	GAATCACTGAGAAATATTTCACTATTATTAAGAGCTTTAAAAAGCATTAAGCTTCAATGTT	780
Db	296464	GAATGACTGAGAAATATTTCTGATTTAGGTGATTTAAAAAGCATTAAGAGCTTCAAGAGTA	296405
QY	781	GTTTATAGTGTTATGATGACAGACAAAGCATTTTGAACAAGATTAACGTTGTCAGGA	840
Db	296404	GTATTATGATGTTATGATGATGCTGACAAAGGTATTAATTTGAACAAGATTAACGTTGATCGGC	296345
QY	841	TATGACACATGAACAGTAAAGCACTCGATGTCGTAAATTAATTTGGATCTCTGCGAA	900
Db	296344	TATGACACATGAGGAAGGTAAAGCTATTTGTCATTTGATTAATTAATTTGGATCTCAAGTTGA	296285
QY	901	AAAGTACTTAAACAGCATAGAGAAATTTGAGATGATGATCACTAAAGATTTCAATTTTAA	960
Db	296284	AAAGTACTTAAAGCAATGAAAAAATTTCACTGATGATGTTAGAAATGAAATTTCAATTTTAA	296225
QY	961	GATTATGACAAATGCTTTTGTGTGCTGCTAAAGAACGCACAAAGATTACGTACATTATTC	1020
Db	296224	GATTATGCTCAATTTGCGTCTGCTATCAGCAAAAAGGCGCTAATTAACCATTAATTC	296165
QY	1021	CCTTACATTAATGAAGCAAGTCAAAACCATTAATAAACGTTTCAAAAGTTCAACTTAAAT	1080
Db	296164	CCTTATATCAATCAAGCCAGTGAATAATATAAAAAGCGTGTCCAAAGTCTCACTCAATAT	296105
QY	1081	GAAAGCTGTACTGATGCAATTTCCATGAACCCCTTACACCAAGACAAAGATGAGCGTTTG	1140
Db	296104	GAAAGTGTACTGATGATCCATCTCTATGTAATCCAAACACTTACTGACAAAGGTGAAGACTT	296045
QY	1141	AATGCTTTTATGCAACAACAAGTGTCTATAGAACCAAGCAGACTTTGTTGTTATTTGTTAAT	1200
Db	296044	AATGATTTCTATACAACTCAGTGTCAATGAAACCAAGCAGACTTTGTTGTTATTTGTTCAAT	295985
QY	1201	GATGTAGATTAATGCAATTTTCTTATATAACGCTATTATGAGATTCAAATCCGTCGCGCT	1260
Db	295984	GATGTGATTAATTAATCAATTTTCTTATATAGAGATATTTTGAAGAAATCAAAATAGCTATGCT	295925
QY	1261	TTTGGTTTGAAGGTACACCAATTCATATTTATAGCTGCAAAAGAGA	1305
Db	295924	TTTGGTTTGAAGGAAACACTATTCATATTTATTTCCAAAGAAAGAGA	295880

RESULT 8

AEOI7002 304680 bp DNA linear BCT_16-MAY-2003
LOCUS AEOI7002

DEFINITION Bacillus cereus ATCC 14579 section 5 of 18 of the complete genome.
ACCESSION AEOI7002 AE016877
VERSION AE017002.1 FJ:29894935
KEYWORDS .
SOURCE Bacillus cereus ATCC 14579
ORGANISM Bacillus cereus ATCC 14579
Bacteria; Firmicutes; Bacilliales; Bacilliaceae; Bacillus; Bacillus
cereus group.

REFERENCE 1 (bases 1 to 304680)
AUTHORS Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B., Kapatal,V., Bhattacharyya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonsteln,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
TITLE Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis
JOURNAL Nature 423 (6935), 87-91 (2003)

MEDLINE 22608415
PUBMED 12721630

REFERENCE 2 (bases 1 to 304680)
AUTHORS Candelon,B., Gaillon,X., Ehrlich,D.S. and Sorokin,A.
TITLE The number of ribosomal RNA operons in Bacillus cereus unpublished
JOURNAL 3 (bases 1 to 304680)
AUTHORS Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B., Kapatal,V., Bhattacharyya,A., Reznik,G., Mkhailova,N., Lapidus,A., Chu.L., Mazur.M., Goltsman.E., Larsen.N., D'Souza.M., Walunas.T., Greckin.Y., Fusch.G., Haseikorn.R., Fonslein.M., Ehrlich,D.S.D., Overbeck,R. and Kyripides,N.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Vilvert, Jouy en Josas 78352, France

TITLE JOURNAL FEATURES source

gene cds gene cds gene cds

cds

gene cds

1..304680 /organism="Bacillus cereus ATCC 14579"
/mol_type="genomic DNA"
/strain="ATCC 14579"
/db_xref="ATCC:14579"
/db_xref="taxon:226900"

68..427 /locus_tag="BC1239"
.427 /codon_start=1
/transl_table=1
/product="hypothetical Membrane Spanning Protein"
protein_id="AAP08224.1"
/translation="MLRRKLFLVLPVLAWLISMTKIDMALPLGVETFSNLRIRIQEQLICLFCSIELEFIVSILNLDELFIHRSFTLMKMWNIIQAICYIVEYVALKEIKGOTVFQDNKRME"

551..2041 /locus_tag="BC1240"
551..2041 /locus_tag="BC1240"
/codon_start=1
/transl_table=1
/product="D-lactate permease"
protein_id="AAP08225.1"
/db_xref="GI:29894937"

/translation="MIALPIIMIFLCFLFKOTSLNSSLSIAVSVCIGYLSPFOIGISTHAHTKGWLICEVGVIPEFGELFLPMRAMEGDVARPLEVDRLLOMLMGCIOPLESYSWGCIGFWVAIPFSLSGKPQOAVLSFIGLASWGAMATGTIGSOILNPFLTGLTNTALSIIMEFAVEVLIASHVGGMQVIERKGEVGFLLSEIGSYSNAPVTTELAGIASITVITEGGFLRIKKGSQNITTEAAATERVSTIKIISPYLELVCLLSLRYPALNDPRSVAVLDKSXYKTLLSPGMKMCFTTIERFRPAINTKOSSLQTIKOMDFATTTFWIHSISELGASGHSHSLAKTAGDTCTFFPVAFPSNIATSIGELTGNSAGSNAMPKIQQMQAONVALPWQVVTLTNASSAVTIACC

gene
2132..2818
/locus_tag="BC1241"
2132..2818
/locus_tag="BC1241"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08226.1"
/db_xref="GI:29894938"
/translation="MTLKNYGVLKGTIYQSEIGCKTPHVOHLQGETGVDIRIAIN
KSOSEYSEVLFASNNIKSEAIHILPLPGETTEIKNNPKVALDYLRGNLFDNOMI
PLPEKAGADNLEKIERIKRAIEKAIYAGERMGPEENTPDSYFHRKPNNGIH
DHIMQONVERKQDDGIMODGILLIHFKEKEEWIGIFLAFOSQSMCTDEGHARVP
EHCYKGNIK"
complement(2830..3663)
/locus_tag="BC1242"
complement(2830..3663)
/locus_tag="BC1242"
/codon_start=1
/transl_table=11
/product="integral membrane protein"
/protein_id="AAP08227.1"
/db_xref="GI:29894939"
/translation="MEDVNSIFQILLTSVLIQSLGVIIVIGFLGLGESLFRMYS
RARGRKGFLLTAMIGVPIHELGHAIMCWLFRKIVATOFPTDSOGALGVQYVNO
KSVYORIGNEFIIIGITLILMRVFPRESYELFPTDSTASTINTEMIO
NMILSTFVLKSLPTIGNLNPSFLFLAICISAHIALSKPIKSIDQVYIMFIV
LPLFNITTGLPDYDSNQLIGVMKYNMILAFSSVALLFSCITLVSFSFYKIRGRNS
F"
gene
3860..4630
/locus_tag="BC1243"
3860..4630
/locus_tag="BC1243"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08228.1"
/db_xref="GI:29894940"
/translation="MKPSCHTGNAARAKPCGNGCHSTEEVAVSSGQEEPEAKA
RNETVQAKRFASGIVQFRRHALQATPAIMKSGAIEIRKIVSLVLCIGACIFR
MMSASAVNTRFARDISTPFEGESIVFEFFLLITLVEGFIIVSGKMMKSSFSLE
VFGIIMGTIATPIVILVLSFLSFLFPLILSVATYIGISITVAVLKLDNGDLD
LVYLLIANVLIGIATFVLMVSYSTIIQALTKGTGF"
4650..5885
/locus_tag="BC1244"
4650..5885
/locus_tag="BC1244"
/codon_start=1
/transl_table=11
/product="IG hypothetical 16680"
/protein_id="AAP08229.1"
/db_xref="GI:29894941"
/translation="MNVCTKCGAQFEDGVFCQNGSKRGPRVVKNNKSGTKVGITL
LALFVIVIGLYLGASVYQVAVDMRIITLQERDEKLAIEVTADPSVIMRESL
TLPFSYIKENPSYVNELEKILROGEKRDGIERDPSLTKDGKFFLEDRKLAKTY
YTTLTNEKVSLEKMGKEIDTKDKKEKQYGLPENGVOFOSDYKNDYKLSREK
VLMKOSQVNTIDITLOGOYITVQTNAFGATLYVNOKPVTAAGCEPTTGTGPAVDS
ATTLEBNGESGRETTOYERTITLSSYNLPRQKSTEXTVYVNTPTPATCHVNGEI
FPBSDIRKLTSADLTLSKEQLKTAHNEIYARHGHITQTKMQAVFAKQSWYRENPYF
KGTLDIETNYIELIKSRE"
complement(5931..6362)
/locus_tag="BC1245"
complement(5931..6362)
/locus_tag="BC1245"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08230.1"
/db_xref="GI:29894942"
/translation="MCAGILAQLPDEQERKPKPAYSLQSEFINATYISARAWLSH
DPTFADIDIKATOPASASHAGKSLTFPVGSKITINFSDDTITVTASPSDHALKIVPP
HSEALTSGLSHIOASTSGSSKVTFLFHIFSRDPGAP"

gene
6560..7894
/locus_tag="BC1246"
6560..7894
/locus_tag="BC1246"
/EC_number="1.6.99.3"
/EC_start=1
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase"
/protein_id="AAP08231.1"
/db_xref="GI:29894943"
/translation="MNYVITIGDAAGSATNQIYTNDEANVTLLEGEIYSVQOGL
PYVLSGALASTEKLIAHNVKTRDXYGIDAKVRIEYKVDTEKMKVVAEHRKTDVE
FPYRILATVVRVMEGEGRELOGVILKTIPDAERIILKTLETKNVEDVTIIGGA
IGLEMAEFVLELGGKVRRIERNHDIGITVDAMDYEIKHEDAKHIELITMENAKF
GNERVEAIEDTKGKVKADLYVSGVGNPDTLGGTNLRKHKAGIEVNAVYQTNVD
VYAGDCATRHVYIKELHDHILPLCTANKORSLGLNMLDRRAFKGLTGCIIRFND
LTLARTGLNEKAEAGLHPYTKVAVDSVTNMGVYPMAPLKLILRSDTQQLGGVO
IGEEGVDRIDVIAALFNKMSIHDLBDVDLSVAPYNSVMDPIQQAARRAE"
8055..8465
/locus_tag="BC1247"
8055..8465
/locus_tag="BC1247"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP08232.1"
/db_xref="GI:29894944"
/translation="MAINMKTIEMEIASNARHEDRGHVEEKEVCVGDNATLLI
TKNVCFGRKVEVEFFQDHVIVIQEKDEVEIKRYDALITNSNLKTNKNTLELK
FANGOSINLSDNNGYTKMFLFARQIKSIFKLI"
complement(8496..10037)
/locus_tag="BC1248"
complement(8496..10037)
/locus_tag="BC1248"
/EC_number="1.6.99.3"
/EC_start=1
/codon_start=1
/transl_table=11
/product="NADH oxidase (NOXASE)"
/protein_id="AAP08233.1"
/db_xref="GI:29894945"
/translation="WQLTFTLHHLHFSIDEITMPHEVDKADLPYKLYRNVPTPLS
LEIPLSNNSSTPTPLNEIGIYLMWSEVVOQLQDLSGRNVLRSISPGALVYTMEL
IYKIDIDYPDGIIYVDAANHLRILILREGNFQSYLADALGNRCNHDGGAFTSMFV
KNFRKYNNEFYRLOGDGVYIGOLLECAKQFGYNGVYFOLRALHHLGLSSEG
SVYAVIPSTSEPTNMFPHNDVYHKNKTYSHSLLOIPILAHEHVRSHNDVPMIRK
INEASMIESTNHFOTLAHEEYQYHENAVALPKYERLSYDLOLCRRKYSDDAPFTL
KMDAEELATLLEASLSLPSNDLDGKITENAVSLGKTYNADLNGAVTYNSKT
HSIQPLRYGDLRYPLQSSMTDNNVLFQVSLCHLVGKKDYTNALGRYRIRHOMEA
GILVHKLVFAATYAMGMGHPPLSFDTNSCDQLYGIDAGNETSLIOVPVGAARRNMUK
GALHN"
complement(10053..11987)
/locus_tag="BC1249"
complement(10053..11987)
/locus_tag="BC1249"
/codon_start=1
/transl_table=11
/product="hypothetical Cytosolic Protein"
/protein_id="AAP08234.1"

Query Match 56.2% Score 733.8; DB 1; Length 304680;
Best Local Similarity 72.6% Pred. No. 5.3e-118;
Matches 948; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

DB 244018 ATGCCGAACCAAGTAATAGCATAGTAGTCGCCGAAGCAATGACTATTTTC 244077
61 AATAGAATAGTTGGACAGCTGTTTCGATTGTCAGATAGCGCAGTGTAAACAGAGAT 120
DB 244078 AATAGAATTTGCGAAGAAAGCTTCATCGTAGAAGATATTTCCAGGTGTAAACCGCGCAT 244137
121 CGTATTTATTTCTTACAGTGAGTGTAAACATGATTTCAATATTTATTTGATACAGGTGCT 180
|||||

Db 244138 CGTATTTATAGCGCTGGAGAAATGTTAAATCATGAATTTAATATTATGATACGGGCGGA 244197
Qy 181 ATTGAATTTGGTATGATGACCATTTCCAAACAAATTTAGACGCGAGCAAAATCCGCATTA 240
Db 244198 ATTGATATTTGAGACGAGCCGCTTTTGGACAAATTTGTCACAAAGCTGAACTAGCAAAAT 244257
Qy 241 GATGAAGCGGATGATATTTATTTTATGTTAACGTCGCGTGAAGATGACACAAAGCGAT 300
Db 244258 GATGAAGCGGATGATATTTATTTTATGACAAAGCGTCGAGCGGTGAACACACAGCTGAT 244317
Qy 301 GAAATGTCGCTCAATTTTATACAAATCTAAAAAACCGCGTCTATTAGCGGTAAACA 360
Db 244318 GAAAGAAGTTGCAAAATTTTATATGCTTCAATTAACCGCGTGTACTTGCATTAATAA 244377
Qy 361 GTAGATTAATGAGAAATGCGTACAGACGCTGATGATTTCTATTCATTTAGCAATTTGGTGA 420
Db 244378 GTAGCAATATCCGGAATGCGCGAGTATGATTTATGACTTTTATGACTTTAGCAATTTGGCGAA 244437
Qy 421 CCGTATCCAAATTCAGGGTCAACATGTTAGCTCTGGTGAATTTGTTAGATGACATTTGTT 480
Db 244438 CCATTTCCCAATTTGAGTACACACGATTTAGGGTATGAGTACTTACTACACAGCTGCA 244497
Qy 481 TCTCATTTTGGTGAAGAGAGAGATCCCTTATGATGAAGATACAAATTTGACTATTCAT 540
Db 244498 CAACATTTTCCAAAGATTTGAGAGACGCGATTTGATGAAGACACAAATTTCTTTCTTTA 244557
Qy 541 ATTGACGACCAAAACAGTAAATCAAGTTTAAATGCTATTTTAAAGTGAAGATGCGC 600
Db 244558 ATTGACGCTCCAAAGCTAAGGGAATCATCATCTTGAATATGCTCTCTCTGTCAGAGCGCT 244617
Qy 601 GTTATCGCTTCTTAATGTTGACAGGACAGAGACGCTATGATACAGATATGATAT 660
Db 244618 GTAAATGTTAGTAAATGACGAGAAACACTGCTGATGCTGCTGAACACATATAGTAA 244677
Qy 661 GATGACAAAGATTTATGTTTAACTGATCTGCTGATGCTGATGCTGATGCTGATGCTGAT 720
Db 244678 GATGAAAAAGATTTATGATATTTTATGATACGCTGCTGATGCTGCTGATGCTGAT 244737
Qy 721 GATCACTGAGAAATTTCTACTATTAAGAGCTTTTAAAGCGATTTGAACGTTCAATGTT 780
Db 244738 GAAATGACAGAAATATATAGTACTGCTGCGCTAAGAGCGATTTGAACGTTCTGAGT 244797
Qy 781 GTTTAGTGTATTTGATGACGAGAACAGGCAATTTGAACAGATTAACGTTTTCAGAGA 840
Db 244798 GTTTAGTGTGTATTTGATGAGTGAAGAGAAATTTATGAGCAAGATTAACGTTTTCAGAGA 244857
Qy 841 TATGACATGAGACAGTAAAGCAGTGTGATGCTGATTAATTAATGATGCTGATGAGAA 900
Db 244858 TATGCTCATGATTCAGGCGAGCTGTTGTTATCGTTGTTAATTAATGGATGCTGATGAA 244917
Qy 901 AAAGATAGTAAAGCAGTAAAGAAATTTGAAGATGAAGTAAAGATTTCCAAATTTTAA 960
Db 244918 AAAGAGCAAAAAACAATGAAACATTTGAAAGAAATATCCGCTCATTTCCAAATTTCTTA 244977
Qy 961 GATTATGACAAATTTGCTTTTGTCTGCTAAAGAACGCAAGATTACGTTACATTTAT 1020
Db 244978 GAATAGCGACCAATTTGCTTTCTTATCTGCAAAAACGAGAAAAAGTAAACAAACATTAATA 245037
Qy 1021 CCTTACATTAATGAAGCAAGTAAACCAATTAACCAAGTCAAGTTCACTTTAAAT 1080
Db 245038 CCGGTTATCGATGAGAGTAAAGTAAAGTATCCGTTATTCAAACGAACTATTTAAAC 245097
Qy 1081 GAAATGCTTACTGATGCAATTTTCATGAACCTTACACCAACAGCAAGATGAGAGCTTTG 1140
Db 245098 GATGAATTAATGATGCGGTATCGATGAATTCGACGCAACATCATATGATGATGCGTGG 245157
Qy 1141 AATGCTTTTATGCAACACAGTGTCTATAGAACCGACGACATTTGTTGTTATTTAT 1200
Db 245158 AAAATTTTCTATGCAACACAGTGTGCGTAAACACGACATTTGTTGTTGTTAAAC 245217
Qy 1201 GATGAGATTAATGATTTTCTTATTAAGCACTTATTAAGAGAAATGATGCTGCGCT 1260
Db 245218 GATCTGAAATTAATTCACCTTCTCATATAGAGCTTTCTTAAGAAATGCGCTTACGTAATCA 245277.

Qy 1261 TTGGTTTTGAGGTACACCAATTCATATTAATGCTCGAAGAGA 1305
Db 245278 TTGGCTTCGTAGGAACGCGGATTCACATTAATGCTAGAGCAAGA 245322
RESULT 9
AE017028
AE017028
LOCUS
DEFINITION
Bacillus anthracis str. Ames section 5 of 18 of the complete genome.
ACCESSION
AE017028
AE016879
VERSION
AE017028.1
KEYWORDS
GI:30255149
SOURCE
Bacillus anthracis str. Ames
Bacillus anthracis str. Ames
ORGANISM
Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
1 (bases 1 to 290117)
Read,T., Peterson,S., Tourasse,N., Ballille,L., Paulsen,I., Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Oksstad,O., Helgason,E., Ralston,J., Wu,M., Kolonay,J., Beaman,M., Dodson,R., Brinkac,L., Gwinn,M., Deboy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H., Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weldman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.
The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria
NATURE 423 (6935), 81-86 (2003)
MEDLINE
22608414
PUBMED
12721629
TITLE
2 (bases 1 to 290117)
Read,T., Peterson,S., Tourasse,N., Ballille,L., Paulsen,I., Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Oksstad,O., Helgason,E., Ralston,J., Wu,M., Kolonay,J., Beaman,M., Dodson,R., Brinkac,L., Gwinn,M., Deboy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H., Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weldman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.
Direct Submission
Submitted (26-Mar-2003) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
FEATURES
source
1..290117
/organism="Bacillus anthracis str. Ames"
/mol_type="genomic DNA"
/strain="Ames"
/db_xref="taxon:198094"
331..1929
/locus_tag="Ba1197"
331..1929
/locus_tag="Ba1197"
/locus_tag="Ba1197"
/note="similar to GP:10717143, and GP:10717143; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="oligopeptide ABC transporter, oligopeptide-binding protein"
/protein_id="AAP25159.1"
/db_xref="GI:30255150"
/translation="MMLACSQKNEPQANAKGDSKSGAKOVINLTPVSEIPMDT LASDAASSKVMNMEGLYLRLGDKLPLVGAKSHTKSEDKRTFKLRDEKMSNGD EYTAQDFVYAMQRAVNPPTAKSGYIMLDVNAKEIKINRELSPDQAGKALIDYTFEV ELNDVPEYFLDLTVPLPEFLPELNEKYMKGQGEKFGLEATTTLYNGFVLEKMHQSPQ LKKNSTYDQKEVKLEEVNFNVKDTGAVNVYNTSALDRGLSLELVKDKGQDFK TINDPVEFLRPKNQALANKNTRKAIKSAAYDRGIEVILNNSKAYGIVPSDFV KQPDQRREKNGKLSKYDVKEAKKFWENAGKEKELGKDIETIEFLNFDNEBKRIGIEY

KGELEKNIPGLTVITITKQPNFAQKIAKLEDSCGFQISFPGMGPDFADATEELDMNTTDS
 OKMKQTSIAEVDILIRKAKQSDVKMGKWMNDLAKKILFEADAIAFVQGRSILIER
 ETIKQVYIHKHYGGELSEFKMASVEK"
 complement(1956. .2099)
 /locus_tag="Bai198"
 complement(1956. .2099)
 /locus_tag="Bai198"
 /note="Identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAP25160.1"
 /db_xref="GI:30255151"
 /translation="MIGVITNSIGEKLFEGCKEMKFVNKSLDVGWNEVGGKGIYKK
 RRY"
 2731. .3126
 /locus_tag="Bai200"
 2731. .3126
 /locus_tag="Bai200"
 2731. .3126
 /locus_tag="Bai200"
 /note="similar to SP:P32687, PID:409796, GB:U00096,
 PID:1790459, SP:P32687, PID:409796, GB:U00096, and
 PID:1790459; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAP25161.1"
 /db_xref="GI:30255152"
 /translation="MVTLYSSPSCSCGRKATLMEENHIPTEINISDPDLTEEIKE
 ILRMESTDEILISRSKVPQELVWNESELPDQLYKINRYPQILRRPIMIDEKRLQ
 VQYNDEIRRLPRVTRFQRLREARLYN"
 complement(3176. .3850)
 /locus_tag="Bai201"
 complement(3176. .3850)
 /locus_tag="Bai201"
 complement(3176. .3850)
 /locus_tag="Bai201"
 /note="identified by match to PFAM protein family HMM
 PF03772."
 /codon_start=1
 /transl_table=11
 /product="membrane protein, putative"
 /protein_id="AAP25162.1"
 /db_xref="GI:30255153"
 /translation="MDEFLTSLYALMGVIDVYDAGDNDAIYALASRLPESKRKAIL
 IGTIAIVRILITLAVYLLDIPPLQIGSVLLTLAVNLLTDSNDLSSIGKTTTL
 EQAVTIVFAQVWGFENVVIAAGAAHGRFSLVIGLISIPITIMGSKILIMEMERF
 PFLICGAAIIAYTAGKMYTHEDRLATFFHNNPSTFAGIPYLFITLICGIVQVRR
 LRNVKH"
 4112. .4210
 /locus_tag="Bai202"
 4112. .4210
 /locus_tag="Bai202"
 /note="Identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAP25163.1"
 /db_xref="GI:30255154"
 /translation="MICMSESLHNNKQGRVFWILKELMIIR"
 4207. .4860
 /gene="meca"
 /locus_tag="Bai203"
 4207. .4860
 /gene="meca"
 /locus_tag="Bai203"
 /note="similar to GB:X03350, GB:X03350, GB:X03350,
 GB:DD00137, GB:M21692, GB:M24317, GB:M24308, GB:M24309,
 GB:M24310, GB:M24311, GB:M24312, GB:M24313, GB:M24314,
 GB:M24315, GB:M24316, GB:M37067, GB:X15447, GB:X15448,
 GB:X15449, GB:X15450, GB:X15451, GB:X15452, GB:X15453,
 GB:X15454, GB:X15455, SP:P00325, SP:P00326, SP:P07327,
 PID:1223739, PID:178098, PID:178100, PID:178111,
 PID:199428, PID:28416, PID:598143, PID:825622, GB:X03350,
 GB:X03350, GB:X03350, GB:DD00137, GB:M21692, GB:M24317,
 GB:M24308, GB:M24309, GB:M24310, GB:M24311, GB:M24312,

Query Match 55.7%: Score 727.4: DB 1: Length 290117;
Best Local Similarity 72.3%: Pred. No. 6,9e-117;
Matches 944: Conservative 0; Mismatches 361; Indels 0; Gaps 0;

1 ATGACTAAACCTATAGTACTATTGAGTAGGACCTAAATGAGTAGTAACTACAAATTTT 60
||| ||||| || ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 280374 ATGCCGAAACGATATAGCATATAGTAGGCCGCCGACAGTAGGGAATCTACTATTTT 28043333
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

61 AATGAGATAGTTGGAGAACGTTGTCGATTTGTTGAAGATACGCCAGGTGTAACAGAGAT 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 280434 AATAGATTGTTGGAGAGAGATTTCATTCGTAAGAAAGATATTCAGGTGAACGCCGAT 280493
 QY 121 CGTATTATCTTCAGGTAGTGGTTAAGACATGATTTCAATATTAATTAGAGGTGT 180
 Db 280494 CGTATTATAGGCGCTGAGAAATGCTTAATCATGATTAATCAATTTATGATACAGGTGA 280553
 QY 181 ATTTGAATGGTGATGACCATTCACAACAAATTTAGACGCGACAGCAAAATGCCATA 240
 Db 280554 ATTTGATTTGGAGACGACCATTTTGGACACAATTCGTCACAAACAGCTGAAGTAGCAATT 280613
 QY 241 GATGAGCGGATGTTTATTTATTTATGTTTAACGTGCGAGCATTTGACACAAACGAT 300
 Db 280614 GATGAGACGATGTTATTCATTTTATGACAAATGCTGACGCTGAACAGCGCGAT 280673
 QY 301 GAAATGTCCTCAATTTTATCAATATCTAAACAAACCGCTGCTATTAGCGGTAAACAA 360
 Db 280674 GAAAGAGTTCGCAAAATTTATTTATTCGTTCAAAATAAACCACTTCTACTGAGTAATAA 280733
 QY 361 GTAGATTAATGAGAAATGCGTACAGACGCTGATGATTTCTATTCAATTAGGATTTGTGA 420
 Db 280734 GTGGACAAATCCGGAATGCGCAGTATGATTTGATTCATAGGATTTGGCGAG 280793
 QY 421 CCGTATCCAAATTCAGGCTACATGCTTTAGGCTCTGTCGACCTTGTAGATGCACTTGT 480
 Db 280794 CCATTCCTCAATCTCAGGTACACACGATTAAGATGATGATGATGATGATGATGATG 280853
 QY 481 TCTCATTTTGGTGAAGAGAGAGATCCCTTATGATGAAGATACAAATTCACATATCCAT 540
 Db 280854 CAGCATTTTCCAAAGATTTGAAGAGACGATGATGATGATGATGATGATGATGATGATG 280913
 QY 541 ATTTGACGACCAACATAGTAAATCAAGTTTATGTAATGCTATTTTATGATGATGATG 600
 Db 280914 ATTTGCGCTGCAAAATTTAGGAAATCATCTTGAATGCTCTCTGCTGCAAGACGT 280973
 QY 601 GTTATGCTTCTATGTTGAGGAGACAGAGACGATTCATGATACAGATTTATGAT 660
 Db 280974 GTTATGCTTCTATGTTGAGGAGACAGACGATTCATGATACAGATTTATGAT 281033
 QY 661 GATGACCAAGATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 Db 281034 GATGACCAAGATTTATTTATGATGATGATGATGATGATGATGATGATGATGATG 281093
 QY 721 GATGACCAAGATTTATTTATGATGATGATGATGATGATGATGATGATGATGATG 780
 Db 281094 GAAAGTACAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 281153
 QY 781 GTTATGCTTCTATGTTGAGGAGACAGAGACGATTCATGATACAGATTTATGAT 840
 Db 281154 GTTATGCTTCTATGTTGAGGAGACAGAGATTTATGATACAGATTTATGAT 281213
 QY 841 TATGACATGAACAGATTAAGAGATGATGATGATGATGATGATGATGATGATGATG 900
 Db 281214 TATGACATGAACAGATTAAGAGATGATGATGATGATGATGATGATGATGATG 281273
 QY 901 AAAAGATGTAAGACATGTAAGAAATTTGAAGATGTAAGATTTGAAGATTTGAAG 960
 Db 281274 AAAAGATGTAAGACATGTAAGAAATTTGAAGATTTGAAGATTTGAAGATTTGA 281333
 QY 961 GATTATGACAAATTTCTTTGTTGCTGCTAAAGAGACGACATTTAGCTACATTTAT 1020
 Db 281334 GATTATGACAAATTTCTTTGTTGCTGCTAAAGAGACGACATTTAGCTACATTTAT 281393
 QY 1021 CCTTACATTTATGACAAAGTGAAGAAATTTGAAGATTTGAAGATTTGAAGATTTGA 1080
 Db 281394 CCTTACATTTATGACAAAGTGAAGAAATTTGAAGATTTGAAGATTTGAAGATTTGA 281453
 QY 1081 GAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 Db 281454 GAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 281513
 QY 1141 AATGCTTTTATGACAAAGTGAAGAAATTTGAAGATTTGAAGATTTGAAGATTTGA 1200
 Db 281514 AATGCTTTTATGACAAAGTGAAGAAATTTGAAGATTTGAAGATTTGAAGATTTGA 281573

QY 1201 GATGATGAATTAATGATTTTCTTATTAACAGCTATTTAGAGATCAAAATCCGCGCT 1260
 Db 281574 GATCCAGAACTATTTGACCTCTCATATAGACGCGTCTTATAAATCGTTAGCTGAATCA 281633
 QY 1261 TTTGCTTTTGAAGCTACACCAATTCATTTATGATGATGATGATGATGATGATG 1305
 Db 281634 TTTGCTTTTGAAGCTACACCAATTCATTTATGATGATGATGATGATGATG 281678
 RESULT 10
 AL591981/c 347050 bp DNA linear BCF 06-JUN-2002
 LOCUS Listeria monocytogenes strain EGD, complete genome, segment 9/12.
 DEFINITION
 ACCESSION AL591981
 VERSION AL591981.1 GI:16411141
 KEYWORDS
 SOURCE Listeria monocytogenes
 ORGANISM Listeria monocytogenes
 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 REFERENCE
 1 Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baguer, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T., Chardil, A., Chetouani, F., Couve, E., de Daruvar, A., Deloux, P., Domand, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Enlihan, K.D., Esili, H., Portillo, F.G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Haut, J., Jackson, D., Jones, L.M., Kaerst, U., Kretz, J., Kuhn, M., Kunst, F., Kurapkhat, G., Madueno, E., Maltounam, A., Vicente, J.M., Ng, E., Nedjari, H., Nordstok, G., Novella, S., de Padlos, B., Perez-Diaz, J.C., Pierrel, R., Rimmel, B., Rose, M., Schlutener, T., Sliemers, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehlund, J. and Cossart, P. Comparative genomics of *Listeria* species Science 294 (5543), 849-852 (2001)
 JOURNAL MEDLINE 2153729
 PUBMED 11679669
 REFERENCE 2 (bases 1 to 347050)
 AUTHORS Glaser, P., Frangeul, L. and Rusniok, C.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE
 COMMENT E-mail: pglaser@pasteur.fr
 FEATURES
 source
 gene
 CDS
 /organism="Listeria monocytogenes"
 /mol_type="genomic DNA"
 /strain="EGD-e"
 /db_xref="taxon:1639"
 complement(95, .841)
 /gene="lmo168"
 complement(95, .841)
 /gene="lmo168"
 /note="similar to glucose 1-dehydrogenase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC9766.1"
 /db_xref="GI:16411142"
 /db_xref="SPTREMBL:O8Y636"
 /translation="MNKVALVTGSSSGRLGRLTAIALAKGQYDIAVFSRRNRKAEKVQ OEIEQGRKRCVIFKANVDGKRELFKAVDEDFGRDIFINNAAGVLRPLMELEES HMDTMNINAKALLFAGQBAKLMQHRQSGKISLSISIRLENYTVGVSKAAVE SLRWYLAELAPGIVANVSGLIETDNLNHPNREELLKDAVSKTPAGRMIEPNDL VNAVYFLASEKADIMRGITLVNGRTILY"
 complement(join(845, .1942,1948, .1953))
 /gene="lmo1689"
 complement(845, .1942)
 /note="similar to A/G-specific adenine glycosylase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC9767.1"

[illegible]

/db_xref="SPTREMBL:O8Y6J1"
 /translation="MKTTSISVQDKNREYNIFIDEKYNFSVDEEVLARYQLMKGKVL
 /TEADIEIKQADVNRKGLNAINFLSHRVRSEKIRYLRKOEEMPAIDELIKLKN
 MDYINDEPAEILYKTQIKITTKLKPRTIEBELYKGLTRELISRIEYIEQDAOLENA
 TKQAKIMKRNNSAKMKLOOKITITDILQGGTSELSKAAATEATSTLDAVDEAILQ
 KQIEKTRKKNKRYKPSIAKOKTITSLMQGFSYDTIOSYLTENISFEEB"
 gene
 complement(3872..3877)
 /gene="lmo1692"
 RBS
 complement(3872..3877)
 /gene="lmo1692"
 RBS
 complement(4680..4685)
 /gene="lmo1693"
 join(4757..5673,5681..5708)
 /gene="lmo1694"
 4757..4762
 /gene="lmo1694"
 4771..5673
 /gene="lmo1694"
 /note="similar to CDP-abequose synthase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99772.1"
 /db_xref="GI:16411148"
 /db_xref="SPTREMBL:O8Y6J0"
 /translation="MHILGATGFIGDHLVHLEKSDHELYILTRQKLRKANVHYT
 EMLLEDKLPNLEDLPVDCINILAGAGLMDKWTYDRKRYVNSRIEATSAISIVKMK
 KSKRLIMNSAIGATYVSSKSTIYLQTEENTYADNPLGKTYVEEKTASASDGIEN
 VYANQGLCTGNSGSPFVPEKLTQTYGRGRGNROWYSNIHVDVVAALIFEDHRO
 INGVNFTAPHPOVEKKFAERLQKKMKPKTPPKKIIKFIILGERAMTILDSORATP
 EKLMSNHEFRFEFLQBALDDLD"
 5681..5708
 /gene="lmo1694"
 complement(join(5682..8291,8299..8304))
 /gene="lmo1695"
 terminator
 complement(5682..5709)
 /gene="lmo1695"
 CDS
 complement(5694..8291)
 /gene="lmo1695"
 /note="similar to putative membrane proteins"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC99773.1"
 /db_xref="GI:16411149"
 /db_xref="SPTREMBL:O8Y6I9"
 /translation="MKREKMOAYAMPQKNSYVIVPITPMARVFEIILATGIDY
 PSLKENTSSQPEQIFIMETVGLAATPMILYDVYIKLPGKSPSHVIASIGATNM
 PTNIGFGGVGLGASLRASFYTKNSHSEILIAISKIALDFVSGSLCYLSVALTLIP
 GFADHFNVPWMLLAGGLYEPILFTITKMSKSLFVLDPIKRELTLIASLLEWGFAR
 GCAIFIGLGMPEWDILKEPPLFVIAVIGIASWPGGVGFVVMILGSLQSDOD
 LALAMLFVRFERYIITPEVVGLEFVQKAGKVDLEGJLDFLOKVAHFELVIFVY
 GSGILLSSNVAIYHVPETIKMPENLFTSQITIVAGELLGLARICEKTKKK
 AYITIVILGCAITNLTARVFSMKQALFGLIVLCLFLARNEFREKLYTMSKVIID
 SIIFIEVLAGIYVIGIYNSPMIKHSKEIPDLRIASHELMVLVGVGFIAVSALVIT
 IYLSYTEKLGSPFEAVKVRHKLAKGNEVSHMFLRDLKLFMAAGEVLSFYRIIA
 DKVMINDEPTGNMDKMAALEVYMMNADRQYRPFVEYERVTMTPLVLDHGFDFIKGLG
 EEGVDVQNFMSGKKKKGERALNDRREGYTEIIEPPNHDHTMTTLRAVSDSEMIL
 GREKGFSLGFEFDYVLEQAPIALAKGCEGYVGFASMPSPYTDQMSIOLMRSKEEA
 PSGIMDFLITNLFERAKEDGQFOTRNAGMAPLANYGESKLYAFGLRGLAGLYVRSQGY
 GFKLNRFSKRYVTWEQKFAFRKRSSIAFTMLQMLIVGKKRPLANSOVVDLPLE
 EETKKPDSE"
 RBS
 complement(8299..8304)
 /gene="lmo1695"
 gene
 complement(8318..9007)
 /gene="lmo1696"
 CDS
 complement(8318..8992)
 /gene="lmo1696"
 Query Match 52.4%: Score 683.2; DB 1; Length 347050;
 Best Local Similarity 70.2%: Pred. No. 2.9e-109;
 Matches 916; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

```
D 259438 ATGGCAAAACACGTTGACGATTTGTGGAGCTCCAAACGTTGGCAAAATGCATATTTT 259379
Q 61 AATGAATAGTTGAGAACGTTTTCGATTTGTGAAAGATACGCCAGGTGAACAGCAT 120
D 259378 AACGATCGTTGGTGAACGTTGTTCATAGTGAAGATGTTCCGGGTGACACGATGAC 259319
Q 121 CGATTTATTTCTCAGGTGAGTGGTTAACACATGATTTCAATATTTATTTGATACAGTGT 180
D 259318 CGCATATATATTCAGCGGAATGCTTGAAGAAACAAATTTAACAATATTTGATACAGTGT 259259
Q 181 ATTGAATTTGTTGATGACGATTCACAAACAAATTTAGACGCGGAGGAGAAATGCCATA 240
D 259258 ATTGATTTTTCGATGACGATTTTGTAGACAAATTTCCGACACAGCGGAATGACAAAT 259199
Q 241 GATGAACGGAGTTTATTTATTTTATGTTTAAAGTGGTGAAGGATTTGACACAAAGCAT 300
D 259198 GATGAACGAGCATTAATTTATTTTATTTTAAACAAATGTCGTGAAGAGTTACGATGACAC 259139
Q 301 GAATGTCCTCAATTTTATACAAATCTAATAAACCCTGCTATTAGCGGTTAACAAA 360
D 259138 GAACAAATGACAAAATTTCTTACCGGCTAATTAACCAATTTGTTTACGATTAATAAAA 259079
Q 361 GTGATATATATGAAATGCGTACAGACGATGATGATTTCTATTCATTAAGATTTGGTGA 420
D 259078 GTGATATACCCAGAAATGCCAGATCAAAATTTATGACTTTTATCTTGGTTGGTGGAG 259019
Q 421 CCGATTCATATTCAGGTCACATGTTTAAAGTCTGTCGTGATGATTTAGATGACATGTT 480
D 259018 CCGATTCATATTTCTGTTACATGATGACGCTTGGCAGATGCTTGAATGCTGTTGCC 258959
Q 481 TGTATTTTGTGTAAGGAAAGAAAGATCCTTATGATGATGATGATGATGATGATGATGAT 540
D 258958 GGTCAATTTTCCAAAAGAAAGAAAGAAATTCACATGACACAGAAATTCAGTTTA 258899
Q 541 ATTGAACGACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
D 258898 ATGCGTGACCAATGTTGTTAATCTTCTATTTCTAATGACATGCTTGGAGAAATGCT 258839
Q 601 GTTATCTTTCTAATGTTGACAGGACACAGACAGCCATTTGATACAGATATGTTAT 660
D 258838 GTTATGTTTTCAGATATTTGGGGTACAACTGCTGATCAATTTGATCACTTATACCTT 258779
Q 661 GATGACAAATGATGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 720
D 258778 GATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 258719
Q 721 GAATCAATGAAATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
D 258718 GAAGACACAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258659
Q 781 GTTTTGTGTTTATGATGACAGAAACGATTTATGACAGATTAACGATGATGATGATG 840
D 258658 GTTCTGTGTTTATCAACGAGAAAGATGATGATGATGATGATGATGATGATGATG 258599
Q 841 TATGACATGAAACGATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 900
D 258598 TATGCGCATATGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 258539
Q 901 AAGATAGTAAACGATGAAATTTGAAGATGATGATGATGATGATGATGATGATGATGAT 960
D 258538 AAGATGAAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 258479
Q 961 GATTTATGACAAATGCTTTTGTGCTGTTAAAGACGACAGATTAACGATGATGATGAT 1020
D 258478 AGCTATATGACCAATTTGTTTGTATGTTGTTAAACAAACAGCTTAATAATCTATTC 258419
Q 1021 CCTTACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
D 258418 CCACTTATTAACCAAGTAACGATTAACGATTAACGATTAACGATTAACGATTAACGAT 258359
Q 1081 GAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
D 258358 GATGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 258299
```

```
Q 1141 AATGCTTTTATGCAACAAAGTTGCTATAGAACACCGACATTTGTTGTTTAAAT 1200
D 258298 AAGATTTTCTATACACAAAGTGGCGTAAACACCAACGTTGTTGTTGTTTAAAT 258239
Q 1201 GATGATGATTAATGATGATTTTCTTATTAACCGTATTTAGAGATCAAAATCCGCGCT 1260
D 258238 GATCAGAACATCAATGATTTTCTTATGATGATGATGATGATGATGATGATGATGAT 258179
Q 1261 TTTGCTTTTGAAGTACACCAATTCATATTTATGATGATGATGATGATGATGATGAT 1304
D 258178 TTTCAATTTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 258135
```

```
RESULT 11
AX641670/c 349980 bp DNA linear PAT 21-FEB-2003
LOCUS Sequence 2860 from Patent WO0101118.
DEFINITION AX641670
ACCESSION AX641670
VERSION AX641670.1 GI:28474431
KEYWORDS Listeria monocytogenes
SOURCE Listeria monocytogenes
ORGANISM Listeria monocytogenes
REFERENCE 1 Bacterla; Firmicutes; Bacillales; Listeriaceae; Listeria.
```

```
AUTHORS Buchrieser, C., Frangoul, L., Couve, E., Rusnok, C., Fsihl, H.,
Denoux, P., Dussurget, O., Chetoui, F., Nedjati, H., Glaeser, P.,
Kunst, F., Cossart, P., Daniels, J., Geobel, W., Kreft, J., Kuhn, M.,
Ng, E., Vassez, Boland, J., Dominguez-Bernal, G., Garrido-Garcia, P.,
Tierrez-Martinez, A., Amend, A., Chakraborty, T., Domann, E., Hahn, T.,
Berche, P., Charbit, A., Durand, L., Perez-Diaz, J. C., Baquero, F.,
Garcia del Portillo, F., Gomez-Lopez, N., Medueno, E., de Pablos, B.,
Wehlund, J., Kaerst, U., Entian, K. D., Hain, J., Rose, M., and Voss, H.
Listeria monocytogenes genome, polyepitides and uses
Patent: WO 0101118-A 2860 11-Apr-2001;
INSTRUT PASEUR (FR)
```

FEATURES

source location/Qualifiers

1. 349980

/organism="Listeria monocytogenes"

/mol_type="genomic DNA"

/db_xref="taxon:1639"

/note="Seq 1 for big: original length: 2944528-splitted into 10 <23>seq 0001 : From 0.000.001 to 0.349.980 length: 349.980 <23>seq 2855 : From 0.300.001 to 0.649.980 length: 349.980 <23>seq 2856 : From 0.600.001 to 0.949.980 length: 349.980 <23>seq 2857 : From 0.900.001 to 1.249.980 length: 349.980 <23>seq 2858 : From 1.200.001 to 1.549.980 length: 349.980 <23>seq 2859 : From 1.500.001 to 1.849.980 length: 349.980 <23>seq 2860 : From 1.800.001 to 2.149.980 length: 349.980 <23>seq 2861 : From 2.100.001 to 2.449.980 length: 349.980 <23>seq 2862 : From 2.400.001 to 2.749.980 length: 349.980 <23>seq 2863 : From 2.700.001 to 2.944.528 length: 244.528"

BASE COUNT 99761 a 75437 c 61177 g 113605 t

ORIGIN

Query Match 52.4%; Score 683.2; DB 6; Length 349980;

Best Local Similarity 70.2%; Pred. No. 2.9e-109;

Matches 916; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

```
Q 1 ATGACTTAACCTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
D 212438 ATGCAAAACAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 212379
Q 61 AATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
D 212378 AACGAAATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 212319
Q 121 CGTATTAATTTCTCAGGTGAGTGAATTAACGATGATGATGATGATGATGATGATGATG 180
D 212318 CGCATATATTAATTCAGCGGAATGCTTGAAGAAATTTTAAATTTATGATGATGATG 212259
```

Oy	181	ATTGAATGTGATGACCAATTC	CAACAAATTTAGAGCGCAGCAAAATGCGCATTA	240
Db	212258	ATTGATTTTCCGATGAACATTTT	TATAGCAAAATTCGACCAAGCGAAATTCGCAAT	212199
Oy	241	GATGAAGCGGATGTTATTTATTT	TATAGTAAAGTGGTGAAGATTCACCAAGCAT	300
Db	212198	GATGAAGCAGACGTAATTTATTT	TATACCAATGGCTGTGAAGGATTAACCGATCAAC	212139
Oy	301	GAATGCGCTCAATTTTATTTAT	TATACCAATTTAAAAACCGGTCGTTAGCGGTTAA	360
Db	212138	GAACAGTAGCAAAAAATTCCTT	TACCGGCTTAATTAACCAATTTGTTTACCGATTAA	212079
Oy	361	GTAAGATAATGAAATGCGTAC	AGACGCTGATGATTTCTATTCATTTAGATTTGGTAA	420
Db	212078	GTAAGATAATGAAATGCGTAC	AGACGCTGATGATTTCTATTCATTTAGATTTGGTAA	212019
Oy	421	CCGATCCAAATATCAGGTCAC	ATGTTAGTCTTGGTACTTTTATGATTCAGATTTGTT	480
Db	212018	CCGATCCAAATTTCTGTTTCA	ATGACTAGGCGCTGGCAGACATCTTGATGCTGTCGC	211959
Oy	481	TCTCATTTTGGTGAAGAGAA	AGATCCTTATGATGAAGATACAAATTCGACTATCCATT	540
Db	211958	GCTCATTTTCCAAAAAGAA	AGAGAAATATCCAGATGACACAGTGAATTCAGTTTA	211899
Oy	541	ATTGACGACCAAAACGTAGT	AAATCAAGTTTATGTAATGCTATTTTATGAGTGAATCGC	600
Db	211898	ATCGGTGACCAAAATGTTGT	TAATAATCTCTATTTATTAATGCACTGCTTGAGAAATCGT	211839
Oy	601	GTTATCGTTTCTAATGTTGC	AGGCAACAGCGTATTTGATACAGATATAGTTAT	660
Db	211838	GTTATGTTTTCAGATATTG	CGGTGACCACTGATGCAATTTGCAATTCATTAATCTTTC	211779
Oy	661	GATGACCAAGATTATGTTT	ATATCATGCTGCTGATGCTGATGCTGATGCTGATGCTGAT	720
Db	211778	GATGACCAAGATTATGTTT	ATATCATGCTGCTGATGCTGATGCTGATGCTGATGCTGAT	211719
Oy	721	GAATCACTGAGAAATATTC	AGTATTAAGAGCTTTAAAGCGATTGAACGTTTCAATGTT	780
Db	211718	GAAGACAGAGAAATATAG	CTTTTACGTGCAATGAGACGAATTAACCGCTCGGACGTT	211659
Oy	781	GTTTATGAGTTTATTTGAT	GACAGAACGCAATTTATGAACAAGATTAACGTTTTCAGAGA	840
Db	211658	GTTTATGAGTTTATTTCA	CGCAGAAAGATTAATCGTGACGAATTAACGATTAATGCTGGA	211599
Oy	841	TATGACATGAAACAAGT	AAAGACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT	900
Db	211598	TATGACATGATGCCGAC	GTGCATTTATTTGTTGTAACAAATTTGGATGCAATTTAAC	211539
Oy	901	AAAGATAGTAAACGATGA	AGAAATTTGAAGATGAGTACGTAAAGAAATTCCAATTTTGA	960
Db	211538	AAAGATGAAAAACAATTA	ATAGTATGACAGAAATTTCTGACCAATTCGCAATTTCTTA	211479
Oy	961	GATTATGACAAATTTGCTT	TGTGTCTCTAAGAACGACACAGATTACGTACATTAATTC	1020
Db	211478	AGCTATGACCAATTTGTT	TGCTATCTAAACAAACGCTTGAATTAATCTATATTC	211419
Oy	1021	CCTTACATTAATGAAGCA	ATGAAACCATTAATAAAACGTTCAAGAGTTCAACTTTAAT	1080
Db	211418	CCACTTATTTAACCAAGT	ATACCATTCATTAAGTATGATGATGATGATGATGATGATGATGAT	211359
Oy	1081	GAAGTGTGTACTGATGCA	ATTTTCCATGAAACCTTACACCAACAGACAAAGGTAGACGTTTG	1140
Db	211358	GATGTTATTTAGTAGAT	GCTGTGCAATGAACCATACCTATGATGATGATGATGATGATGATGAT	211299
Oy	1141	AATGCTTTTATGACACAG	AGTTGCTATAGACACGACATTTGTTGTTATTTGTTAAT	1200
Db	211298	AAGATTTTCTATACACAG	AGTGGGTTAAACCAACCTTTGTTGTTATTTGTTAAT	211239
Oy	1201	GATGAGATTAATGATTTT	CTTATTAACGCTATTTAGAGATCAATCAATCGTCCGCT	1260
Db	211238	GATCAGAACTCATGATTTT	CGTATGACGCTTTCTTATGAAACCGGATTTAGACAGCA	211179
Oy	1261	TTTGGTTTGAAGGTACAC	CAATTCATATTAATAGCTGAAAGAG	1304

Db	211178	TTTCCATTTGAGGTACGCAATTCGAGTAATTTGCTGTAACGC	211135
RESULT 12	AL596170/c		
LOCUS	AL596170	313450 bp	DNA linear BCT 06-JUN-2002
DEFINITION	Listeria innocua Clp11262 complete genome, segment 8/12.		
ACCESSION	AL596170		
VERSION	AL596170.1	GI:16414292	
KEYWORDS			
SOURCE	Listeria innocua		
ORGANISM	Listeria innocua		
REFERENCE	1		
AUTHORS	Glaser, P., Frangoul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloeker, H., Brandt, P., Charraorty, T., Charbil, A., Chetouani, F., Couve, E., de Darvar, A., Denoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fajhl, H., Portillo, F. G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Krefl, J., Kuhn, M., Kunz, F., Kurapkut, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordiek, G., Novella, S., de Pablos, B., Perez-Diaz, J. C., Purcell, R., Rammel, B., Rose, M., Schlueter, T., Simoes, N., Trierer, A., Vazquez-Boland, J. A., Voss, H., Wehlend, J. and Cossart, P.		
TITLE	Comparative genomes of Listeria species		
JOURNAL	21537279		
MEDLINE	21537279		
PUBMED	11679669		
REFERENCE	2 (bases 1 to 313450)		
AUTHORS	Glaser, P., Frangoul, L. and Rusniok, C.		
TITLE	Submitted (09-Jul-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE		
COMMENT	E-mail: pglaser@pasteur.fr		
FEATURES	Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.		
SOURCE	location/Qualifiers		
gene	1..313450		
CDS	/organism="Listeria innocua"		
	/mol_type="genomic DNA"		
	/strain="Clp11262"		
	/db_xref="taxon:1642"		
	complement(287..1444)		
	/gene="lin1790"		
	complement(287..1444)		
	/note="similar to transmembrane transport proteins"		
	/codon_start=1		
	/transl_table=1		
	/protein_id="CAC97021.1"		
	/db_xref="GI:16414293"		
	/db_xref="SPTREMBL:Q92AX8"		
	/translation="MGAGKFWILITLVVAISGLSGVLPLPLAIIIEKGVASAGINGF		
	HATGIYLVIVLSPFIEAPLHKYGRKPIIVVGGLVAIAIFAPIMFNLVFWFIRLL		
	IGIDGHIHFSQWTWIGAMSDPSKGRNMAIVGFESVGAIPQVLNLAKNANLPE		
	FLSGIIVLAWSLVFWIRNDPGEKAVIRKISFMGSIKRSFGEVFMAYAMIPPLYG		
	ILIEGTANTFPVYGLRDGLDMITAMIISSPSVGTITFVPGIISDKFRGRKMLPL		
	TGAGAVVFLVAEVRIPLEFVFEFFLILGLSLVSLGSLYMTDLPLPLFAGNIVL		
	GMFSLGISIGPSATGMIGIFGNQIFVFAVAGILLGLCLLALGKREAV"		
	complement(1508..1579)		
	/product="rRNA-Glu"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 69.21"		
	complement(1585..1675)		
	/product="rRNA-Ser"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 63.76"		
	complement(1712..1784)		
	/product="rRNA-Asn"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 85.46"		
	complement(1791..1864)		
	/product="rRNA-Ile"		
	/note="rRNA-scan-SE vs 1.3 result - Cove score = 92.91"		

QY	121	CGATTTATTTCTCAGGGAGGGGTTAACACATGATTTCAATATTAATTATACAGGTGCT	180
Db	276690	CGCAATTAATTAATTCGCGGAAGGCTTGAAAAAATTTAAACATTAATGATACAGGTGGT	276631
QY	181	ATTGAAATTTGGGATGACACATTTCCAAACACAAATTTAGAGGCGACAGGCAAAATTCGCATTA	240
Db	276630	ATTGATCTTTCCGACGAACCATTTCTTAGAGCAAAATTCGGCACAAAGCGGAATTCGCAATT	276571
QY	241	GATGAAGCGGATGTTATTAATTTTATGTTTAACGTCGCGGAAGGATTTGACACAAAGCGAT	300
Db	276570	GATGAAGGAGAGCGTAATTAATTTATTTATACCATGTCGCGGAAGGGGTTACCGATGCAGAC	276511
QY	301	GAATGTCGCTCCAAATTTTATACAAATCTAAAAAACCGGTCGATTTAGCGGTTAACAA	360
Db	276510	GACCAAGTAGGACAAAATTTCTTACCGGCTTAATAAACCAATTTGTTTACCGATTAATTA	276451
QY	361	GTAGTAAATATGGAATGCGTTACAGCGGTATGATTTCTATATTCATTAGCATTTGGGTGA	420
Db	276450	GTAGATTAACCCAGAAATGCGGTATCAGATTTATGACTTTTATCTTCTGGTTGGGTGAG	276391
QY	421	CCGTAATCAATATCAGGGTCACATGGTTAGGCTCTTGAGACTTGTATAGCATGTTGTT	480
Db	276390	CCGTAATCAATTTCTGGTTCTCATGAGACTAGGGGCTTGGGATATGCTTATGCTGTTCCG	276331
QY	481	TCTCATTTTGGTGAAGAGGAGAGATCCTTATGATGAAGATACAATTTGCATATTCATT	540
Db	276330	GCTCATTTTCCAAAAGAGAGAGAGAGATCCAGATGAAGAACTGTAATTTAGTTTG	276271
QY	541	ATTGGAGCACCAGACTAGTAAATCAACGTTTAGTAATATGCTATTTAGGTGAAGATCCG	600
Db	276270	ATTGGTCGGCCAAATGTTGGTAAATCACTATTTCTTAACGCACTTCTTGGCGAAGCCG	276211
QY	601	GTTATCGTTTCTAATGTTGACAGGCAACGAGAGACGCTATTGATACAGATATAGTAT	660
Db	276210	GTCATGTTTCTGCACATTCGGGGAACCTACTCGTGATGCAATTTGATTCACACTTATACGTTT	276151
QY	661	GATGACACAAGATTATGTTTAAATCGATCTGCTGATGCTGTAATAAAGGAAAAAGTATAT	720
Db	276150	GATGGTCAGAGATTATGTCATGATTTGATACAGCCGGAATGAGAAACGTTGGGAAGGTAT	276091
QY	721	GAATCACTGGAATATTTCACTATTAGAGCTTTAAAAAGCATTAACGTTCAATGTT	780
Db	276090	GAAGCACAGAGAAATATAGTGTTTTACTGCAATGAGAGCAATTAAGCTCCGATGTT	276031
QY	781	GTTTATAGGTTATTTGATGACAGAACGCAATTTATGAACAAGATTAACGTTGAGAGG	840
Db	276030	GTTCTGTTGGTTATCAACGCCAGAAAGGATATTCGAGACCAAGATTAAGCGGATTCGTGGA	275971
QY	841	TATGCACATGAACAAGGTAAAGACAGTCTGATTGTCTGTAATTAATGGATACTGTGAA	900
Db	275970	TATGGCATATGATGCGGACGGCCTATTCATTATTGTATGAACAAATGGGATGCAATTAAC	275911
QY	901	AAAGTTATTAACCATGAGAAATTTGAAGATGAAGTACGTAAAGATTTCCATTTTAA	960
Db	275910	AAAGATGAAAAAAACAATTAATGTATGAGCAGAGAAATATTCGGGAGCAGTTCCAAATCTTA	275851
QY	961	GATTATGACAAATGCTTTTGTCTGCTGAAGAAGCAGCAAGATTTAGCTACATTAATTC	1020
Db	275850	AGCTATGACCAATTTGTTTCTATCTGCTAAAACAAAACACGCTTAAACAACATTAATTC	275791
QY	1021	CCTTACATTAATGAAGCAAGTGAAAAACCAATAAAAAACGTGTTCAAAATTCAACTTTAAAT	1080
Db	275790	CCGCTCATTAACCAAGTAAGCATTAATCATTTATTCATTAACGCGGTACAAATCTAGTATGTA	275731
QY	1081	GAAGTCGTTACGTATTCATTTCCATGAAACCTTACACCAACAGACAAAGGTAGACGTTTG	1140
Db	275730	GATGTTATTTAGATGCGGTTGCAATGAATTCATCCACCAATGGATTAAGGTAAACACACTT	275671
QY	1141	AATGCTTTTATGAACAACAGTGTCTATAGAACCCGCAATTTTGTGATTTGTTAAT	1200
Db	275670	AAAATATTTCTATACACACACAGTGGCTGTTAAACGCCCAACTTTTGTGATTTGTTAAT	275611
QY	1201	GATGTAGATTAATGCAATTTTCTTATATAACCGCTATTTAGAGATCAACAAATCCGTGCCGCT	1260

Db	275610	GATCCAGAACTAATGATTTCTTTATGAACGTTTCTTGAAGAACCGGATTAAGAAGCA	275551
Qy	1261	TTTGGTTTGAAGTACACCAATTCATATATAGTCGAAAG	1304
Db	275550	TTTCCGTTTGATGTCACGCCAATTCCAGTAATTCGTCGTAAGCG	275507
RESULT 13			
AX413016			
LOCUS	AX413016	319630 bp	DNA
DEFINITION	Sequence 7 from Patent WO0228891.		linear
ACCESSION	AX413016		
KEYWORDS	AX413016.1	GI:21445474	
SOURCE			
ORGANISM	Listeria innocua		
REFERENCE	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.		
AUTHORS	Kunst, F. and Glaser, P.		
TITLE	Listeria innocua, genome and applications		
JOURNAL	Patent: WO 0228891-A 7 11-APR-2002; INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR) Location/Qualifiers 1. .319630 /organism="Listeria innocua" /mol_type="genomic DNA" /db_xref="taxon:1642"		
FEATURES			
source			
BASE COUNT	105207 a 53428 c 66726 g 92263 t	6 others	
ORIGIN			
Query Match	51.9%; Score 676.8; DB 6; Length 319630;		
Best Local Similarity	69.9%; Pred. No. 3.8e-108;		
Matches 912; Conservative 0; Mismatches 392; Indels 0; Gaps 0;			
Qy	1	ATGACTTAACCTATAGTACTATTGAGTAGGCGCTAATGTAGTAATCTACAATTTT	60
Db	50618	ATGGCAAAACCAAGTTGACGATTTGCGGACGTGCAAAAGTTGGCAATCGACTATTTT	50677
Qy	61	AATAGAAATGTTGGAGACGCTGTTGCGATTTGGAAGATACGCCAGTGTACACGAGAT	120
Db	50678	AACAGAAATGTTGGTGAACGCTGTTCCATAGTGAAGATGTTCCGGTGAACGCGAC	50737
Qy	121	CGATTTATTTCTTCAGAGTGGTGAACATGATTTCAATATTTATGATACAGNGGT	180
Db	50738	CCCATATATTAATTCACGCCGAATGCTTGGAAAGAAATTTTACATTATGATACAGGTGT	50797
Qy	181	ATTGAATTTGGTATGACACCATTCCAACACAATTAAGACCGCAGCAGAAATCGCAT	240
Db	50798	ATTGATCTTTCGAGACCAACATCTTAAGCAAAATTCGGGCACAAAGGGAAATCGCAAT	50857
Qy	241	GATGAACCGGATGTTATTTATTTTATGTTTAACGTCGCTGAAGGATTTGACACAAGCAT	300
Db	50858	GATGAAGCGAGCGTATTTATTTTATTAACCAATGCTCGAAGGGGTACCGATGCGAC	50917
Qy	301	GAATGGTGGCTCAATTTTATTAACAATTAATAAAAAACCGGCTGATTAAGCGTTAA	360
Db	50918	GAAACAAGTAGCAAAAATTTCTTACCGGCTCAATAAACCATAATTTTATAGCGATTAATA	50977
Qy	361	GTAGATTAATATGGAATGCGTAGACAGACGTATAGATTTTCTATTCAATAGAGATTGGTAA	420
Db	50978	GTAATATTAACCAAAATGCGTGATCGATTTATGACTTTTATTTCTGTGGTTGGTGAG	51037
Qy	421	CCGATTCCAATATCAGGCTACATGGTTTAAAGTCTTGGTGACATTTGATGCAAGTTGT	480
Db	51038	CCGATATCCATTTCTGGTTCTCATGACCTAAGGCGCTTGATGATGCTGATGCTGTCGT	51097
Qy	481	TCTCATTTTGGGAAGAGAAGATCCATTATGATGAAGATCAATTCAGATATCCATT	540
Db	51098	GCTCATTTTCCAAAGAGAAGGGAAGATACCCAGATGAACACAGTGAATTTAGTTTG	51157
Qy	541	ATTGGACGACCAACCTAGTAATCAAGTTTATGTAATTCATTTTATAGTGAAGATCGC	600

Db 51158 ATTGGTCGGCAATGTTGGTAAATCATCTATTCTTAACCCACTCTTGGCGAAGACCG 51217
QY 601 GTTATCGTTTCTAATGTTTGCAGGACAGACAGACGCTATTGATACAGATATTAGTTAT 660
Db 51218 GTCATGTTTCTGACATTTGGGGAACTACTCGTATGATCAATGTATCACTATTAGCTTT 51277
QY 661 GATGACAGATATTATGTTTAAATCGATCTGCTGATGCTGTAAGAAAAGAAAGTATAT 720
Db 51278 GATGGTCAGATTTTATCATGATTTGATACAGCGGATGTGAAGAAACCGTGAAGTGTAT 51337
QY 721 GAATCAACTGAGAAATATTTCAGTATTAGAGCTTTTAAACCGATTGAGCTCAATATGTT 780
Db 51338 GAAGACACAGAGAAATATAGTGTTTTACGTGCAATGAGCAATTTGAAACCTCCGATGTT 51397
QY 781 GTTTAGTGTATTATGATGAGAACAGACGATTATTGAACAGATTAACGTTGTCAGGA 840
Db 51398 GTTCTGTGGTTATCAACGACAGAAAGATATTGAGAGCAATTAAGCGATGCTGGA 51457
QY 841 TATGCACTGACAAAGTAAAGCATGCTGATTTGCTAATTAATGGGATCTGTCGAA 900
Db 51458 TATGCCATGATGCCGAGCGGCTATCATTTATGTTGTAACAAATGGGATGCAATTAAC 51517
QY 901 AAGATAGTAAACAGATGAAGAAATTTGAAGATGAAGTAAAGATTAAGATTCATTTTA 960
Db 51518 AAGATGAAAAAACAATTAATGATGACAGAAAGATATTGGAAGCGATTCATTTTA 51577
QY 961 GATTATGACAAATTCCTTTTGTCTGCTAAAGAACGCAAGATTACGATTAATTC 1020
Db 51578 AGCTATGCAACCAATTTGTTTCTGATCTGTAACAAACAAACAGCTTAACCACTATTTC 51637
QY 1021 CCTTACATTAATGAACAGTGAAGACATTAACAAACGCTTCAAGTTCACTTAAT 1080
Db 51638 CCGCTATTATACCAATATAGCGATTAATCATTTCTTACGCTAACAACTACTAGTCAAT 51697
QY 1081 GAAGTCTTACTGATCAATTTCCATGAACCCCTACACCAAGCAAGTAAAGTAAAGCTTTG 1140
Db 51698 GATGTTATTAATGATGCGGTTGCAATGAATCCATACCAATGATTAAGTAAACGACTT 51757
QY 1141 AATGCTTTTATGCAACAGAGTGGTATAGAACACCGACGATTTGTTATTTGTTAAT 1200
Db 51758 AATATATTCATACACACAGAGTGGCTTAAACCGCAACATTTGTTATTTGTTAAT 51817
QY 1201 GATGATGAATATGATTTTCTTATTAACGCTATTAGAGATCAATCCGTCGCCCT 1260
Db 51818 GATCCGAACATAATGATTTCTTATGAACGTTTCTTGAAGAAACCGGATTAAGAGAAGA 51877
QY 1261 TTTGTTTGAAGTACACCAATTCATTTATATGATGCTGAAGAG 1304
Db 51878 TTTCCGTTTGTATGATGACGCAATTCAGATTAATTTGCTGTAAGCG 51921

RESULT 14
AX417046/c 349980 bp DNA linear PAT 02-SEP-2002
LOCUS AX417046 Sequence 4037 from Patent W00228891.
DEFINITION AX417046
ACCESSION AX417046 GI:21449656
VERSION
KEYWORDS
SOURCE Listeria innocua
ORGANISM Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

REFERENCE
1 Kunst, F. and Glaser, P.
Autors
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 4037 11-APR-2002;
INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES
source Location/Qualifiers
1. 349980
/organism="Listeria innocua"
/mol_type="genomic DNA"
/db_xref="taxon:1642"

/note="seq 2058, original length: 3,011,208 replaced
by-seq 2058: 0.000,001 to 0.349,980-seq 4032: 0.300,001 to
0.649,980-seq 4033: 0.600,001 to 0.949,980-seq 4034:
0.900,001 to 1.249,980-seq 4035: 1.200,001 to
1.549,980-seq 4036: 1.500,001 to 1.849,980-seq 4037:
1.800,001 to 2.149,980-seq 4038: 2.100,001 to
2.449,980-seq 4039: 2.400,001 to 2.749,980-seq 4040:
2.700,001 to 3.049,980-seq 4041: 3.000,001 to 3.011,208"
BASE COUNT 101055 a 72969 c 60688 g 115268 t
ORIGIN

Query Match 51.9% Score 676.8; DB 6; Length 349980;
Best Local Similarity 69.9%; Pred. No. 3,7e-108;
Matches 912; Conservative 0; Mismatches 392; Indels 0; Gaps 0;

QY 1 ATGACTTAAACCTTATAGTATGATTTAGTATGAGCCCTTATAGTAAATCTACAAATTTT 60
Db 266810 ATGCAAAACACAGTTGATGATGATTTGCGACGCTCAACAGTTGGCAATGATATTTT 266751
QY 61 AATAGATAGTGGAGAACGTTTGCATTTGTTGAAGATACGCCAGGTGAACAGAGAT 120
Db 266750 AACAGATCGTTGGTGAACGTTTCCATAGTGAAGATGTTCCCGTGACACGAGAC 266691
QY 121 CGTATTATCTTCAGGTGATGATGATTAACATGATTTCAATTTATGATACAGGTGT 180
Db 266690 CGCATATATATTTACGCGAATGCTTGAAGAAATTTAATTTATGATACAGGTGT 266631
QY 181 ATTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 266630 ATTGATCTTTCCGACGACGATCTTTAGACAAATTTGCGCACAGGGAATTCGCAAT 266571
QY 241 GATGAACGAGATGTTATTTTATTTTATGCTTTAAGCTGCGTGAGAGATTGACACAAAGAT 300
Db 266570 GATGAACGAGATGTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTAT 266511
QY 301 GAATGATGCTCAAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 360
Db 266510 GAACAGATGACAAATTTCTTACCGCTTATTTATTTATTTATTTATTTATTTATTTAT 266451
QY 361 GTAGATATTTGGAATTTGCTATGACAGCTGATGATTTTATTTATTTATTTATTTATTT 420
Db 266450 GTAGATATTTGGAATTTGCTATGACAGCTGATGATTTATTTATTTATTTATTTATTT 266391
QY 421 CCGTATTCATTTATGAGGTCACATGTTTATGCTTTGATGATTTTATTTATTTATTTAT 480
Db 266390 CCGTATTCATTTATGAGGTCACATGTTTATGCTTTGATGATTTTATTTATTTATTT 266331
QY 481 TCTCATTTTGGTGAAGAGAGAAATCCCTTATGATGAAGATTCATTTGATGATTCAT 540
Db 266330 GCTCATTTTCCAAAGAGAGAGAAATCCCAATGATGAAGATTCATTTGATGATTCAT 266271
QY 541 ATTGACGACCAACGATGATTAATCAAGTTTATGATGATTTTATTTATTTATTTATTTAT 600
Db 266270 ATTGACGACCAACGATGATTAATCAAGTTTATGATGATTTTATTTATTTATTTATTT 266211
QY 601 GTTATGTTTCTATGTTTGCAGGACAGACAGACGCTATTATATACAGATTAATGAT 660
Db 266210 GTTATGTTTCTATGTTTGCAGGACAGACAGACGCTATTATATCACTATTATTA 266151
QY 661 GATGACAGATTAATGTTTATCGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 266150 GATGACAGATTAATGTTTATCGATGATGATGATGATGATGATGATGATGATGATGAT 266091
QY 721 GAATCACTGAGAAATATTTCAGTATTAGAGCTTTTAAACGCAATTTGATGATGATGAT 780
Db 266090 GAAGACACAGAGAAATATTAGTATTACGTGCAATGAGACATTTGAACGCTCCGATG 266031
QY 781 GTTTAGTGTATTATGATGACAGACAGCATTATTGAACAGATTAACGATTAACGATG 840
Db 266030 GTTCTGTGTTATCAACGACAGAAAGATTTGATGACAGATTAACGATGATGATG 265971
QY 841 TATGCACTGAACAGTAAAGCATGCTGATGCTGATTAATTAAGGATTAATGATGATG 900


```
Db 265970 TATGCGCATGATGCGGAGCCGCTATCATTTAGTGAACAATGGATGCAATTAAC 265911
QY 901 AAGATAGTAAACGATGAAGAAATTTGAGATGACGATGAAGATTTCCATTTTGA 960
Db 265910 AAGATGAAAAAACATTAATGATGTGACAGAGATATTTGGGAGCGATTTCCATTTCTTA 265851
QY 961 GATTATGACAAATTTGCTTTGTGTCTGCTAAAGACGACAAAGATTACGTACATTAATTC 1020
Db 265850 AGCATGACACCAATTTGTTTGTGATGCTGCTAAACAAACACGCTTAACACATTTATTC 265791
QY 1021 CCTTACTTTAATGACGAAGTGAACCATTAACCAACGCTGTTCAAGTTCACTTAAT 1080
Db 265790 CCGCTCAATTAACCAAGTAACGATTAATCATTAATACGCGTACATGTAGTATCTTAAT 265731
QY 1081 GAAGTGTACTGATGCAATTTCCATGACCCCTACACGACGACAAAGATGACGTTTG 1140
Db 265730 GATGTTATTTAGTATGCGGTTTGCATTAATCCATCCACAAATGATTAAGTAAGACGCTT 265671
QY 1141 AATGCTTTTATGACACAAAGTTGCTATAGAACACGACGACATTTGTTATTTGTTAAT 1200
Db 265670 AAATATTTCTATACAAACAAAGTGCGCTTAACCGGACACATTTGTTATTTGTTAAT 265611
QY 1201 GATGTAGATTAATGCAATTTTCTTATTAACGCTATTTAGAGATCAATCCGCGCT 1260
Db 265610 GATCCAGAACTAATGCAATTTCTTATGAACGTTTCTTAGAAACCGGATTAAGAGACA 265551
QY 1261 TTTGGTTTGAAGTACCAATTCATATTATAGCTGGAAGAG 1304
Db 265550 TTTCCGTTTATGATGATACGCCAATTCGAGTAAATTTGCTGTAGCG 265507

RESULT 15
AX144037 960 bp DNA linear PAT 31-MAY-2001
LOCUS AX144037
DEFINITION Sequence 2759 from Patent WO0134809.
VERSION AX144037
KEYWORDS AX144037.1 GI:14282890
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Kimmerly, W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 2759 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
SOURCE Location/Qualifiers
1..960
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 336 a 131 c 189 g 304 t
ORIGIN
Query Match 51.2%; Score 667.6; DB 6; Length 960;
Best Local Similarity 81.2%; Pred. No. 6.8e-106;
Matches 775; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
```

```
Db 181 CTTTCTATCATCGGTAGACCTAATGTTGTAATCTAGCTTGTCAATGCTATTTTACGC 240
QY 592 GAACATGCGCTTATCGTTTCTTAATGTTGACAGGACAAACGAGACGCTATTGATACAGAG 651
Db 241 GAAGACGCTGTTATGTTGCTTAATGTTGCTGTAACCTGACAGATGCGCATTTGATACGAG 300
QY 652 TATAGTTATGATGACAAAGATTAATGTTTAAATCGATACGCTGATGATGATTAAGAAAGA 711
Db 301 TACTCTTATGATGACAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 712 AAGATATATGATCACTGACGAATATTCAGTATTAAGAGCTTTAAAGCAATGGAACGT 771
Db 361 AAGGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 772 TCAATGTTGTTTATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
Db 421 TCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 832 GTTCAGATATGACATGACAAAGATTAACGACGCTGATGATGATGATGATGATGATGATGATGAT 891
Db 481 GTACGTGCTATGACATGACAAAGATTAACGCTATGATGATGATGATGATGATGATGATGATGAT 540
QY 892 ACTGTGAAAAAGATAGTAAACGATGAAGAAATTTGAAGATGAAGTACGTAAGAAATTC 951
Db 541 ACAGTTGAAAAAGATAGTAAACGATGAAGAAATTTCACTGATGATGATGATGATGATGATGAT 600
QY 952 CAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
Db 601 CAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 1012 ACATTAATTCCTTACATTAATGAAAGCAAGTGAAGAAACATTAAGAAACGTTCAAGTTCA 1071
Db 661 ACATTAATTCCTTATATCAATCAACGCGTGAAGAAATCAATTAAGAAACGTTCAAGTTCT 720
QY 1072 ACTTTAAATGAAGTCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
Db 721 ACCTAAATGAAGTCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1132 AGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
Db 781 AGAAGACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1192 TTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
Db 841 TTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1252 CGTGCCGCTTTTGTGTTGAAGTACACCAATTCATTAATTAAGCTGGAAGAGA 1305
Db 901 CGTATGCTTTTGTGTTGAAGAAACACCTATTAATTAATTAATTAATTAATTAATTAATTAAT 954
```

Search completed: September 12, 2003, 19:27:32
Job time : 5331.98 secs

THIS PAGE BLANK (USPTO)

ACCESSION	uncultured marine virus genomic clone SI051pD7L, genomic survey sequence.
VERSION	AY079759
KEYWORDS	AY079759.1 GI:24744504 GSS.
SOURCE	uncultured marine virus
ORGANISM	uncultured marine virus
REFERENCE	Viruses: environmental samples.
AUTHORS	1 (bases 1 to 745)
TITLE	Breibart,M., Salamon,P., Andresen,B., Mahafy,J.M., Segall,A.M., Mead,D., Azam,F. and Rohwer,F.
JOURNAL	Genomic analysis of uncultured marine viral communities
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
PUBMED	222964988
COMMENT	12384570 Contact: Rohwer F Biology Dept. San Diego State University 5500 Campanile Dr, San Diego, CA 92102, USA Tel: 6195941336 Fax: 619595676 Email: forest@sunstroke.sdsu.edu Class: shotgun. Location/Qualifiers 1..745 /organism="uncultured marine virus" /mol_type="genomic DNA" /db_xref="taxon:186617" /clone="SI051pD7L" /clone_lib="Scripts Pier (La Jolla, CA) uncultured virus community" /note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)"
BASE COUNT	167 a 221 c 226 g 127 t 4 others
ORIGIN	
Query Match	11.5% Score 44.4; DB 28; Length 745;
Best Local Similarity	48.7%; Pred. No. 0.34;
Matches	114; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
OY	86 CGGTTCAACAATCTTATGAATCAATACAGCTGTGACCATTCCTATTATGC 145
Db	745 CGGTTCCCAAGCGCAGANAAATAGCTTTCCAGATGCTCTTTTGCGCTTCCCTTC 686
OY	146 TACTTTGTTAACCCTATATACGACGGTTTTTAAGTTTGATATAAATTGAGGACCAT 205
Db	685 TGCCNTGTTGCCAACGACGACACGAGGCACTTCCTGCTCGCGCAACGCGGCAATTCG 626
OY	206 TTCAATCGCTTTGTGTTCAATCCTTCACGACGCTTAACCTAAATAATATAACGCTTC 265
Db	625 NTCCTCATCGGTGTGTAATGCCAGACGCGCATGCATAGCAACAATCGCGGCTCGCATTC 566
OY	266 ATCTATGCGCATTTTCTCGCTGCGCTCTTAATTTGTGTGTTGAATGGTGCAATCAC 319
Db	565 GTCAATGCTGACTGCTGGTCTGCTGCGCATGCTCGCGGAGGAGGCGTCATCC 512
RESULT 8	
CNS000D1/c	
LOCUS	1101 bp DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TRF3 end of BAC #
ACCESSION	BACR01J16 of RPECI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION	AL065414 GI:4938827

	KEYWORDS	GSS.
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACS. For further information please see http://www.fruitfly.org/The_BDP/Drosophila_melanogaster_BAC_library.html Aarón Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BCSPC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.	
FEATURES	location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone_1="BACR01J16" /clone_1lb="RPCI-98" /note="end : TET3"	
BASE COUNT	280 a 104 c 123 g 211 t 383 others	
ORIGIN		
Query Match	10.7%; Score 41.6; DB 29; Length 1101;	
Best Local Similarity	13.8%; Pred.No.2.1;	
Matches	34; Conservative 128; Mismatches 84; Indels 0; Gaps 0;	
OY	138 ATATTATGCTTTGTATTAAAGCGGTAATGAAGCAAGGGTTTTTGAATGTTGATAAATTGGA 197 ::: ::: : : : : : : : : : : : : : : : : : : Db 1099 MTNNIMUUTMMNINNTNMIMTMNMTMTMTTNNTTNNTMNTMTNTMTTUSMU 104d ::: ::: : : : : : : : : : : : : : : : : : OY 198 GGAGCATTCGACTGCCTTGCTGCTCACGCCAGCCGTAACAGTAAAGAATAGACA 257 ::: ::: : : : : : : : : : : : : : : : : : Db 1039 MSNNNNHMAAMMSMTUTTWTUMMNOMNSSTUUUYUUYUUTNAANNMYNHSNA 980 ::: ::: : : : : : : : : : : : : : : : : : OY 258 TGSCCTCATCATGATGSGAGATTTGCSGSCGCTCATATTGTGTGGAATGSGATCA 317 ::: ::: : : : : : : : : : : : : : : : : : Db 979 TMGNHATNMNMSMINNMNCSTUGRAMUCSMSCMAUMAWMACUYNHCNHTMGNN 920 ::: ::: : : : : : : : : : : : : : : : : : OY 318 CCAATTTCATACASCACTGTATCAATTAATTGAATGATGTGTTAACCAATTCASCTGAAA 377 ::: ::: : : : : : : : : : : : : : : : : : Db 919 WSTATGTGMNHNMMNHNNTNNHHNNHMNTGCUWUWHUUCUWNHYUWUWUWTWT 860 ::: ::: : : : : : : : : : : : : : : : : OY 378 GAATGA 383 ::: ::: Db 859 MHHTAH 854	
RESULT 9	AZ525216 653 bp DNA linear GSS 07-MAY-2001	
LOCUS	AZ525216	
DEFINITION	241PbDev05 Pb MBN #21 Plasmodium berghei genomic 3', genomic survey sequence.	
ACCESSION	AZ525216	
VERSION	AZ525216.1 GI:13965843	
KEYWORDS	GSS.	
SOURCE	Plasmodium berghei	
ORGANISM	Plasmodium berghei	

REFERENCE 1 (bases 1 to 653)
 AUTHORS Carlton, J.M.-R. and Dame, J.B.
 TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
 JOURNAL Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
 COMMENT Contact: Dame JB
 Dept. of Pathobiology, College of Veterinary Medicine
 University of Florida
 2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
 Tel: 352 392 4700
 Fax: 352 392 9704
 Email: damej@mail.ufl.edu
 Seq primer: M13(-20) Forward
 Class: Shotgun.

FEATURES
source

1. 653
 /organism="Plasmodium berghei"
 /mol_type="genomic DNA"
 /strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
 /db_xref="taxon:5821"
 /dev_stage="asexual blood forms"
 /lab_host="Mus musculus"
 /clone_1lb="Pb MBN #21"
 /note="Vector: Bluescript SK(+) vector DNA, phagemid excised from lambda ZAP; Site_1: EcoRV; Site_2: EcoRV; Genomic DNA was prepared from asynchronous blood stage forms of the cloned ANKA isolate of P. berghei grown in laboratory Swiss white mice. The DNA was purified from contaminating host DNA by Hoechst Dye 33258-CSCI ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., ImberSKI, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcoRV-cleaved and dephosphorylated Bluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."

BASE COUNT
ORIGIN

Query Match 10.6%; Score 41.2; DB 28; Length 653;
 Best Local Similarity 51.6%; Pred. No. 2.4;

Matches 94; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 13 CTTCACCAATGAGAAACATGCTAACAAGTCAACCAAGCTTAACCATGACG 72
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 176 CTTCACCTGCATTAAGTCCACATATGAACCAACCACTCGATTATCTTCAATTTT 235
 QY 73 CTGATATCGGATACGGTTCAACCAATCTATGATGAATCAACACGCTGTGACGA 132
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 236 AACAAACAGTTTCGATTAACAAGTACTGATATATCAACATCTTCTTTGTG 295
 QY 133 TTTCATATATCTACTTTGTTAAACCGCTAATACGACGGTTTGTGATATAAAA 192
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 296 CTTTAATATGTTTGTCTGTGAACGCTGATATTTCTAATATTAATTTCTGTAAAA 355
 QY 193 TT 194
 |||||
 Db 356 AT 357

RESULT 10
 LOCUS A2617991 528 bp DNA linear GSS 13-DEC-2000
 DEFINITION IM0449G02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 accession A2617991
 version A2617991.1 GI:11740181
 keywords GSS.
 source Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 528)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah
 Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0449 row: G column: 02
 Seq primer: CACACAGAAACAGCTATGACG
 Class: plasmid ends
 High quality sequence stop: 528.

FEATURES
source

1. 528
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0449G02"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_1lb="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PMD42nv: Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory mouse DNA resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114[9b]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

Query Match 10.5%; Score 40.8; DB 28; Length 528;
 Best Local Similarity 53.0%; Pred. No. 2.9;

Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 101 TAATGAATAGAAATACATACACGCTGTACCATTTATATCTACTTTGTAACCGC 160
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 57 TTATTAATGATAGTACATGCTCGAACAATTTCTTACCTCTTCTTAAGCT 116
 QY 161 TAATGACGCGGTTTATGATTTGTAATAATTGAGCAGCATTTTCATCGCTTGTGT 220
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 117 TATCACATGTTATATGTTGTTTATTAATTAATTAATTTTCTTACCTCTTTCAAG 176
 QY 221 CAATCTTACGCGCATTAACATATAAATAATTAATCAATCCGCTT 264
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 177 AATGATGAAAGCATTTCTTACATTAAGAAAAATATATTTCCGCTT 220

RESULT 11

[illegible]

This clone is available at CHGC in Shanghai.
 Location: Quid44:ora

FEATURES	Location/Qualifiers
source	1. .709

```
/organism="Homo sapiens"  
/mol_type="rna"  
/db_xref="taxon:9606"  
/clone="FHTBAAE05"  
/tissue_type="hypothalamus"  
/dev_stage="Fetal"  
/lab_host="BM25.8"  
/clone_lib="FHTB"
```

	BASE COUNT	ORIGIN
	185 a	149 c
	145 g	227 t
		3 others

	Query Match	10.1%	Score 39.2;	DB 9;	Length 709;
	Best Local Similarity	56.1%;	Pred. No. 8.3;		
	Matches	74;	Conservative	0;	Mismatches 58; Indels 0; Gaps 0
OY	218	TGTCATCTCTTCACGCGATTAACCATTAATAATTAACATCCGCTTCATGTGGCGAT	277		
	247	TGTCATCTCTACATATATTTAGTAGTAGTAACCAACCATACATCTCTTTCTATGCTCTTG	306		
Db					
OY	278	TTTCGCGTGGCGCTTAATTTGTGTTGGAAATGGTCATCACCAAATTCATACCACTCT	337		
	307	TTTCTCCCTTCATATTAATTTTGTCTATAGGACGACCACTGATAGAGTGTGGCANTGCA	366		
Db					
OY	338	ATCATATATATTT	349		
Db	367	AACATTTTATTT	378		

Search completed: September 12, 2003, 20:43:34
Job time : 1049.84 secs

THIS PAGE BLANK (USPTO)